Cosmic Origins Science Enabled by the WFIRST-AFTA Archives

Sally Heap (GSFC) and the SAG8 Team

Carter Chamberlain / Va Tech

Dennis Ebbets / DCEbbets@Gmail.com

Caryl Gronwall/ Penn State

Mario Juric / LSST

Anton Koekemoer / STScI

Jeff Kruk / NASA GSFC

Tom McGlynn / NASA GSFC

Neill Reid / STScI

Chris Stockdale / Marquette U.

Harry Teplitz / IPAC

6 December 2015

TABLE OF CONTENTS

Cosmic Origins Science Enabled by the WFIRST-AFTA Archives

Report by Sally Heap and the SAG8 Team

SUMMARY

The WFIRST-AFTA mission concept is designed to address all three questions in all three of NASA's astrophysics themes: Physics of the Cosmos, Cosmic Origins, and Exoplanet Exploration. This report responds to a charter from Cosmic Origins Program Analysis Group (COPAG) directing COPAG Science Analysis Group #8 to analyze how the WFIRST archive is to be used and to scope the data requirements necessary to conduct science investigations related to the Cosmic Origins theme. It is to describe what data, high-level science data products, and catalogs will be available in the WFIRST-AFTA archive; what data processing is required to produce these data products; and how scientists can identify and access the data of interest.

The report is divided into four sections:

Section 1 provides an introduction to the report.

Section 2 comprises excerpts from the *WFIRST-AFTA 2015 Report* describing Cosmic-Origins science (a.k.a. General Astrophysics) that can be accomplished with data obtained from WFIRST's High-Latitude Surveys (HLS) and the Galactic Bulge Field.

The section also highlights Cosmic Origins science as described by potential Guest Investigators (GI's) using WFIRST-AFTA data and Guest Observers (GO's) who will make new observations of specific proposed targets.

Section 3 briefly describes the data that will be obtained in each of the three High Latitude Surveys (Wide-Area Imaging for Weak Lensing, Wide-Area Spectroscopy for the Galaxy Redshift Survey and Galaxy clustering, the Supernova Field) and Wide-area imaging of the Galactic Bulge. It then goes on to describe the expected processing of data from each of these observing programs and to list sample on-line "catalogs" that will be produced.

Section 4 then describes how WFIRST-AFTA data, high-level science products, and catalogs might be found and accessed through querying the WFIRST-AFTA archive.

SECTION 1: INTRODUCTION

Wide-Field Infrared Survey Telescope (WFIRST) was envisaged by the 2010 Astrophysics Decadal Survey Panel as an "observatory designed to settle essential questions in both exoplanet and dark energy research, and which will advance topics ranging from galaxy evolution to the study of objects within our own galaxy". The transfer of the Astrophysics-Focused Telescope Assets (AFTA) to NASA in 2013 enabled a new, more powerful version of WFIRST, which addresses all three areas of NASA's astrophysics program: Physics of the Cosmos (How does the universe work?) Exoplanet Exploration (Are we alone?), and Cosmic Origins (How did we get here?)

With the submission of 1-page science ideas, it became clear that WFIRST-AFTA observational data obtained for dark energy research or exoplanet demography also could, without modification, address Cosmic-Origins issues. The Final Report of the Science Definition Team (SDT) acknowledges the value of WFIRST-AFTA to Cosmic-Origins research in its Report (http://wfirst.gsfc.nasa.gov/science/sdt_public/WFIRST-AFTA_SDT_Report_150310_Final.pdf) by devoting two sections of its report to general astrophysics. These two sections are reprinted as Section 2 of this report.

In 2014, the Cosmic Origins Program Analysis Group (COPAG) established Science Analysis Group #8 (SAG8) to explore and analyze cosmic-origins science enabled by the WFIRST-AFTA archive. The SAG8 charter is given below:

SAG8 Charter

Cosmic Origins Science Enabled by the WFIRST-AFTA Data Archive

The Wide-Field Infrared Survey Telescope (WFIRST) is the highest priority large space mission recommended by the recent decadal survey in astronomy and astrophysics. It is designed to perform wide-field imaging and slitless spectroscopic surveys of the visible to near-infrared sky. The Astrophysics Focused Telescope Assets (AFTA) study design of the mission makes use of an existing 2.4m telescope to enhance light collecting and imaging performance. The main instrument is a wide-field multi-filter imager with infrared grism spectroscopy. It also features a small-field low-resolution integral field spectrograph. A coronagraph instrument was part of the study and has a primary science focus of direct imaging of gas-giant exoplanets and debris disks.

Achieving the full science potential of WFIRST-AFTA will require input from the astronomical community on how it intends to use the vast WFIRST-AFTA data archive for Cosmic Origins science. The infrared surveys and coronagraphic investigations will provide abundant opportunities for archival research. The high latitude wide-field infrared survey alone is expected to observe more than 500 million galaxies over a 2000 square degree area at a resolution of about 0.11 arc seconds in four broad near-infrared passbands. An active Guest Observer program will further populate the archive with a multitude of datasets. A cross-section of archival Cosmic Origins science investigations would be valuable input in the formative stages of the mission for discussions of high-level science products, catalogs, archive interface design, calibration requirements, data accessibility and distribution, computing resources, and archive operations.

This Science Analysis Group [SAG #8] will analyze how the archive is to be used and scope the data requirements necessary to conduct science investigations related to the Cosmic Origins theme. The SAG will solicit input from the community to identify the types of investigations that will be conducted, and the kinds of data products that are valued and needed. The SAG will also consider what other assets or efforts may be needed to maximize the science return from the WFIRST archive (e.g., coordination of WFIRST-AFTA data with LSST, Euclid, or JWST; GO funding for ground-based observations or theoretical studies). The SAG will document its findings in a report to the Astrophysics Subcommittee.

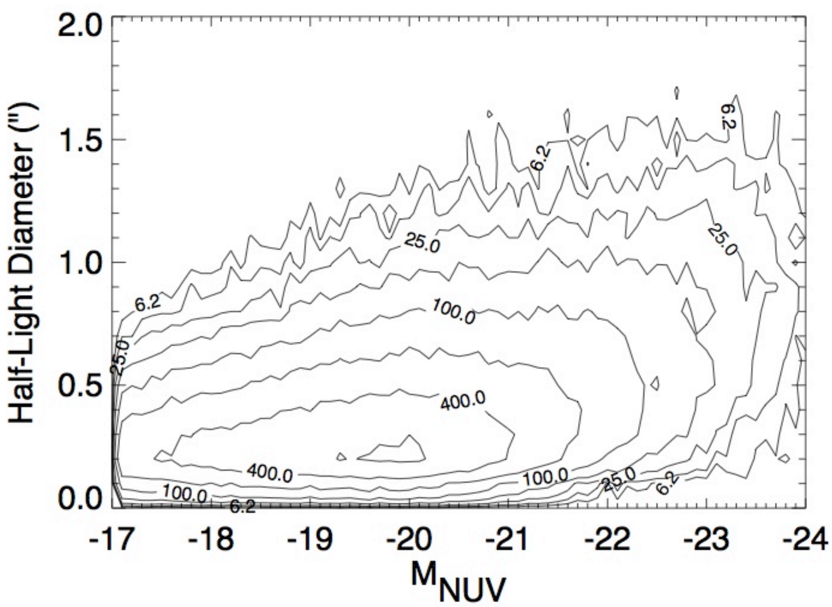
The SAG8 charter can be conveniently divided into three parts. The first part describes and analyzes the WFIRST-AFTA Data Archive from the point of view of Cosmic-Origins scientists – what data, high-level science data products, and catalogs will be available in the WFIRST-AFTA archive; how to access the data; how well the data will be calibrated. The second part describes and analyzes the WFIRST-AFTA Data Archive from the point of view of computer scientists – archive interface design; computing resources; archive operations. The third part describes the WFIRST-AFTA Data Archive from a programmatic point of view – what coordination is needed with archives of other space or ground-based telescopes; funding for Guest Investigators of existing WFIRST-AFTA data, for Guest Observers obtaining new observations with WFIRST-AFTA, and for supporting theoretical studies.

This report comes during an interim period between the WFIRST Science Definition Team (SDT), which concluded its study with a report issued on March 10, 2015 (hereafter referred to as the "Report") and a new Science Working Group whose make-up and responsibilities are as yet unknown. This report therefore only responds to the first part of the SAG8 Charter: an analysis of the WFIRST-AFTA Data Archive from the point of view of cosmic-origins scientists.

A Preview

The COSMOS Mock Catalog (CMC) (Jouvel et al. 2011) gives a preview of the WFIRST-AFTA survey. The CMC is a simulated catalog based on the COSMOS survey of 2 deg^2 by several observatories such as GALEX for UV bands, Subaru for the optical (U to z), and CFHT, UKIRT, and Spitzer for the NIR bands. The CMC contains over 500,000 galaxies, AGN's, and stars. The 111 items in the catalog give information such as type of galaxy, photometric redshift, E(B-V), half-light radius, AB magnitudes in various passbands ranging from Near-UV (Galex) to the IR (Spitzer IRAC). It also gives predicted luminosities of major emission lines in the star-forming galaxies. The data are stored as a structure, so one can extract selected items from the catalog. The figure below gives an idea how the catalog might be used; in this case, to study the frequency of sizes and absolute near-UV magnitudes of galaxies at redshifts, z=0.8-1.4.

Figure 1: Contour plot based on data from the COSMOS Survey



WFIRST-AFTA will survey a high-latitude region 1000 times bigger than the COSMOS field; it will measure the shapes of 400 million galaxies as well as measure magnitudes in 4 passbands; it will obtain grism spectra of the full region. In addition, WFIRST-AFTA will survey the galactic bulge in search of micro-lensing events.

With WFIRST-AFTA's large and varied data sets, it will be impossible to troll a single file for the specific data desired. WFIRST-AFTA data need to be organized differently and stored redundantly for safety. There must be a way to find and access the data you want – and only the data you want -- efficiently.

Organization of This Report

The following sections of this report are organized as follows:

Section 2 comprises excerpts from the *WFIRST-AFTA 2015 Report* describing Cosmic-Origins science (aka General Astrophysics) that can be accomplished with data obtained from the High-Latitude Surveys (HLS) and the Galactic Bulge Field.

The section also highlights Cosmic Origins science as described by potential Guest Investigators (GI's) using WFIRST-AFTA data and Guest Observers (GO's) who will make new observations of specific targets such as the Andromeda galaxy. Only a few examples are given in this report, but Appendix D of the *WFIRST-AFTA 2015 Report* reprints these 1-page science ideas in full.

Section 3 briefly describes the data that will be obtained in each of the three High Latitude Surveys (Wide-Area Imaging for Weak Lensing, Wide-Area Spectroscopy for the Galaxy Redshift Survey and Galaxy clustering, the Supernova Field) and Wide-area imaging of the Galactic Bulge. It then goes on to describe the data processing of each of these types of observing programs and to give a (incomplete) list of on-line "catalogs" that will be produced.

Section 4 then describes how WFIRST-AFTA data, high-level science products, and catalogs can be found and accessed through querying the WFIRST-AFTA archive.

SECTION 2: GENERAL ASTROPHYSICS WITH WFIRST

This section is a reprint of selected portions of the WFIRST SDT 2015 Report, which may be obtained from the WFIRST website:

http://wfirst.gsfc.nasa.gov/science/sdt_public/WFIRST-AFTA_SDT_Report_150310_Final.pdf

or from the arXiv:

Spergel et al., http://arxiv.org/abs/1503.03757v2 .

The first part of this section on pp. 9-15 is a reprint of Section 2.3 of the 2015 Report, which describes general astrophysics with data from WFIRST High-Latitude Surveys.

The second part of this section on pp. 16-19 is a reprint of Section 2.5 of the 2015 Report, which describes general astrophysics with data from the WFIRST Galactic-Bulge Field.

The third part of this section on pp. 20-26 reprints selected Guest Observer and Guest Investigator programs as drawn from Appendix D of the 2015 Report.

As this section is a reprint of a PDF document, the text and its formatting cannot be changed. Thus, the page numbers in the lower right corner are the original page numbers in the 2015 Report. The page numbers in the lower center are the page numbers of this report. In addition, the figure captions with its own numbering system cannot be changed.

Reprinted from the WFIRST 2015 SDT Report (Spergel et al., http://arxiv.org/abs/1503.03757v2)

2.3 High-Latitude Surveys: General Astrophysics

WFIRST-AFTA will produce large data sets with homogenous observing conditions for each set. These data sets will be a treasure trove for Guest Investigators. The HLS will map ~2,200 deg^2 of sky in four broad NIR passbands down to a 5-sigma limiting AB magnitude of J=26.7 and will include a slitless spectroscopic survey component that will obtain R=~600 spectra with a 7-sigma line flux sensitivity of 10^{-16} erg cm^{-2} sec^{-1} over the same region of sky (see Figure 2-15). While the survey is designed to obtain constraints on the dark energy equation of state from weak lensing and BAO measurements, a remarkable range of new astrophysical investigations will be enabled by the HLS data. We focus below on a few of the science cases that are especially well enabled by the access to a 2.4-meter wide-field survey telescope in space and that require no additional time other than that already allocated to the HLS.

2.3.1 *The First Billion Years of Cosmic History*

Mapping the formation of cosmic structure in the first 1 billion years after the Big Bang is essential for achieving a comprehensive understanding of star and galaxy formation. Little is known about this critical epoch when the first galaxies formed. Determining the initial condition and abundances of the primordial interstellar gas as well as the dynamical properties of the first galaxies is critical in order to understand (1) how the first galaxies were formed through a hierarchical merging process; (2) how the chemical elements were generated and redistributed through the galaxies; (3) how the central black holes exerted influence over the galaxy formation and the overall star formation process; and (4) how these objects contributed to the end of the "dark ages."

Finding candidate z > 7 galaxies has been pursued in two ways - from moderate area but very deep imaging surveys (e.g., UDF, Subaru Deep Field, CANDELS) and from images of strongly lensing clusters of galaxies (e.g., Bradley et al. 2008, Postman et al. 2012). Recent observations have identified at least 3 candidate objects at z > 9 (Bouwens et al. 2011, Zheng et al. 2012, Coe et al. 2013). Two of these candidates were found in a survey of massive galaxy clusters. One of the lensed candidates is shown in Figure 2-21. While unlensed sources at these redshifts are expected to be extremely faint, with total AB magnitudes greater than

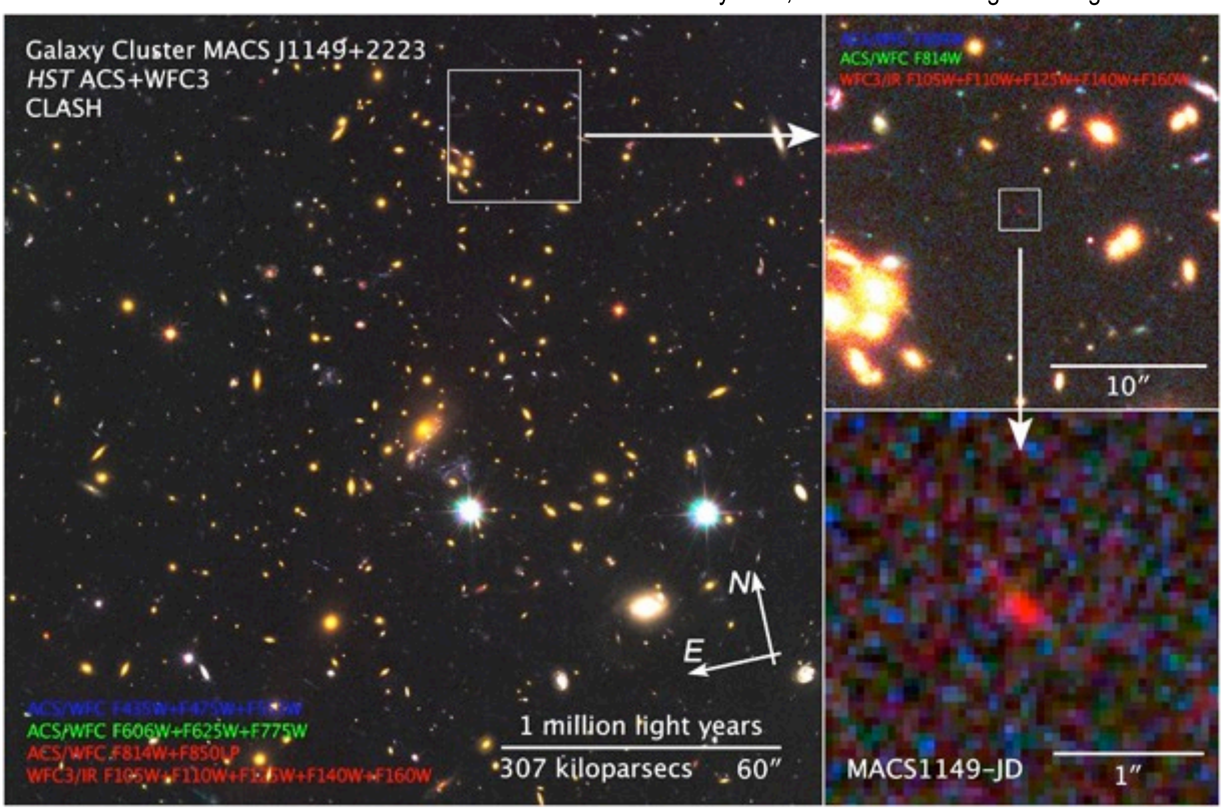


Figure 2-21: A candidate for a galaxy at z=9.6, magnified by a factor of ~15 by the foreground cluster MACS J1149+2223 (z = 0.54). The object was found in an HST survey using the WFC3IR camera (Zheng et al. 2012). This young object is seen when the universe is only about 500 million years old.

28, lensed sources can be 10 – 30 times brighter. This enhancement is particularly important because it puts some of these sources within the reach of the spectrographs on the James Webb Space Telescope and from very large ground-based telescopes. Luminous z > 7 galaxies are extremely valuable as their spectra can be used to determine the amount of Lyα photons produced by stars and galaxies escaping from neutral IGM. These objects will also enable the determination of the time when the epoch of reionization occurred up to its completion. Since only a tiny fraction of neutral hydrogen is needed to produce the high opacity of Lyα observed at z ~ 6, the damped Lyα absorption profile that results from even a partially neutral IGM (Miralda-Escude 1998) can be measured at low-spectral resolution. Furthermore, the early star formation rate (via Lyα and Hα emission measurements; Iye et al. 2006) can be estimated from the spectra of bright high-z galaxies.

The HLS performed by the WFIRST-AFTA telescope as part of its weak lensing program will be superb for finding and studying objects in the early universe. A key advantage that the 2.4-meter aperture provides is that, in the same amount of time, the WFIRST-AFTA HLS will reach ~0.9 magnitudes fainter than the corresponding DRM1 HLS. The final 5-sigma limiting depth of the WFIRST-AFTA HLS is estimated to be J=26.7 AB mag. Figure 2-22 shows the predicted cumulative number of objects that would be found at or higher than a given redshift for both gravitationally-lensed and unlensed regions in a 2,000 square degree survey. The predicted counts (Bradley et al. 2012) for z~8 are extrapolated to higher redshifts using the few known z>9 candidates and bounded by assuming both a pessimistic (dM*/dz = 1.06) and optimistic (dM*/dz = 0.36) evolution of the characteristic magnitude of galaxies. The lensed source count predictions assume ~1 strongly lensing cluster per square degree. The mass models were adopted from the CLASH Multi-cycle treasury program (Postman et al. 2012). The larger collecting area of WFIRST-AFTA relative to DRM1 will yield, in some redshift ranges, as much as 20 times as many z > 8 galaxies (Figure 2-23).

Having 10,000 or more luminous z > 8 galaxies will allow very stringent constraints to be placed on the early star formation rate density, on the amount of ionizing radiation per unit volume, and on the physical proper ties of early galactic structures. With its ~0.1 – 0.2 arcsecond resolution, combined with lens magnifications in the 10 – 30 range, the HLS images will allow us to measure structures on scales of 20 to 50 parsecs, thanks to the boost in spatial resolution provided by the cluster lenses. The HLS will certainly herald a remarkable era in probing the first 1 billion years of cosmic history.

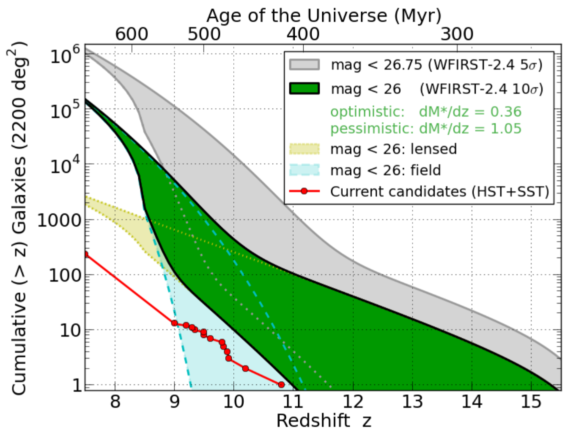


Figure 2-22: Cumulative number of high-z galaxies expected in the HLS. JWST will be able to follow-up on these high z galaxies and make detailed observations of their properties. For understanding the earliest galaxies, the synergy of a wide-field telescope that can discover luminous or highly magnified systems and a large aperture telescope that can characterize them is essential; WFIRST-AFTA and JWST are much more powerful than either one alone.

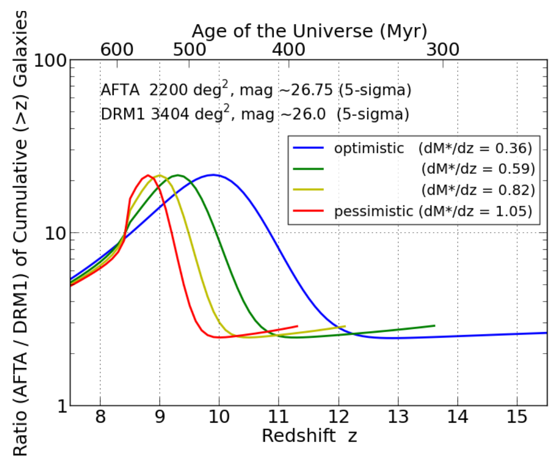


Figure 2-23: The ratio of the cumulative number of high redshift galaxies detected with WFIRST-AFTA to the number detected with a smaller DRM1 version of WFIRST. The 2.4 m aperture yields up to 20 times more high-z galaxies.

2.3.2 *Mapping Dark Matter on Intermediate and Large Scales*

Clusters of galaxies are important tracers of cosmic structure formation. All of their mass components including dark matter, ionized gas and stars are directly or indirectly observable. By its design, the HLS will be superbly suited to mapping weak lensing signatures. The signal-to-noise ratio of the shear signal is proportional to the square root of the surface density on the sky of galaxies that can be used to map the lensing. The WFIRST-AFTA HLS will produce catalogs with surface densities of 60 – 70 galaxies per square arcminute in a single passband and potentially as high as 80 – 100 galaxies per square arcminute in co-added multi-band images. Achieving such densities for weak lensing measurements has already been demonstrated with the WFC3/IR camera on the Hubble Space Telescope. The WFIRST-AFTA HLS will thus enable WL maps that are a factor 3 – 5 higher number density of galaxies than any maps produced from the ground (even with 8 to 10 meter telescopes) as illustrated in Figure 2-24. The larger telescope and multiple bands will produce much more robust and higher quality WL maps than Euclid (see Figure 2-13). On cluster scales (200 kpc – 2 Mpc), this higher galaxy density will allow the dark matter to be mapped to a spatial resolution of ~40 – 50 kpc. When combined with strong lensing interior to clustocentric radii of ~200 kpc, the central dark matter distribution can be mapped down to a resolution of 10 – 25 kpc. Comparing such maps around the hundreds of intermediate redshift clusters expected in the HLS to those from numerically simulated clusters would give us unprecedented insight into the main mechanisms of structure formation.

An especially interesting class of clusters are those in the process of merging / colliding (see Figure 2-25) where all mass components are interacting directly during the creation of cosmic structure. Multiple such mergers have been observed, with the Bullet Cluster being the most prominent example (Clowe et al. 2006). Paired with follow-up numerical simulations, such systems gave important insight into the behavior of the baryonic component (Springel & Farrar 2007) and set upper limits on the dark matter self-interacting cross section (Randall et al. 2008), which is of great importance in the search for the nature of dark matter. Recently, more complicated merging systems have been identified (e.g. Merten et al. 2011, Clowe et al. 2012, Dawson et al. 2012) offering a great opportunity to better characterize the nature of cosmic dark matter. The HLS will be well suited to producing highly accurate strong and weak lensing mass maps in the environs of these information-rich merging systems.

On the largest scales (>10 Mpc), the distribution of matter can best be measured via weak gravitational lensing, owing to the fact that the mass is predominantly dark matter and thus can only be detected gravita-

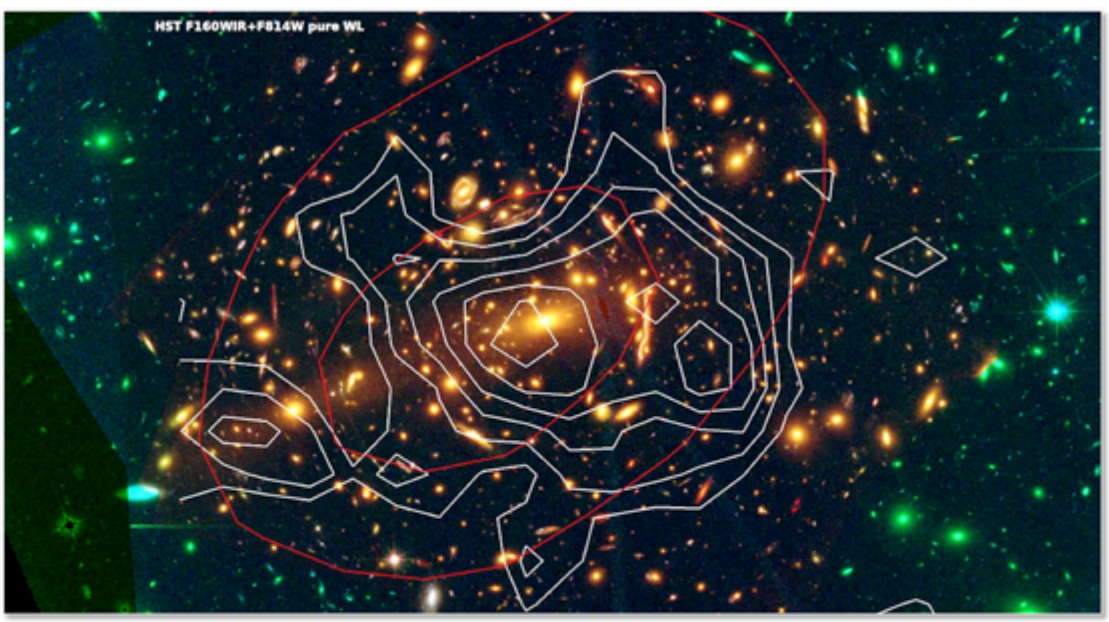


Figure 2-24: Mass density contours around the cluster MACS J1206.2-0848 derived from a ground-based weak lensing survey with Subaru (red) vs. a weak lensing study with HST/ACS+WFC3 (white). The 10x higher surface density of lensed galaxies achieved from space yields ~3x higher spatial resolution maps. The HST data shown here is representative of the WFIRST-AFTA HLS. WFIRST-AFTA will make a map of this quality over 2,000 square degrees as part of its high latitude survey, a thousand-fold increase over the HST COSMOS mass map.

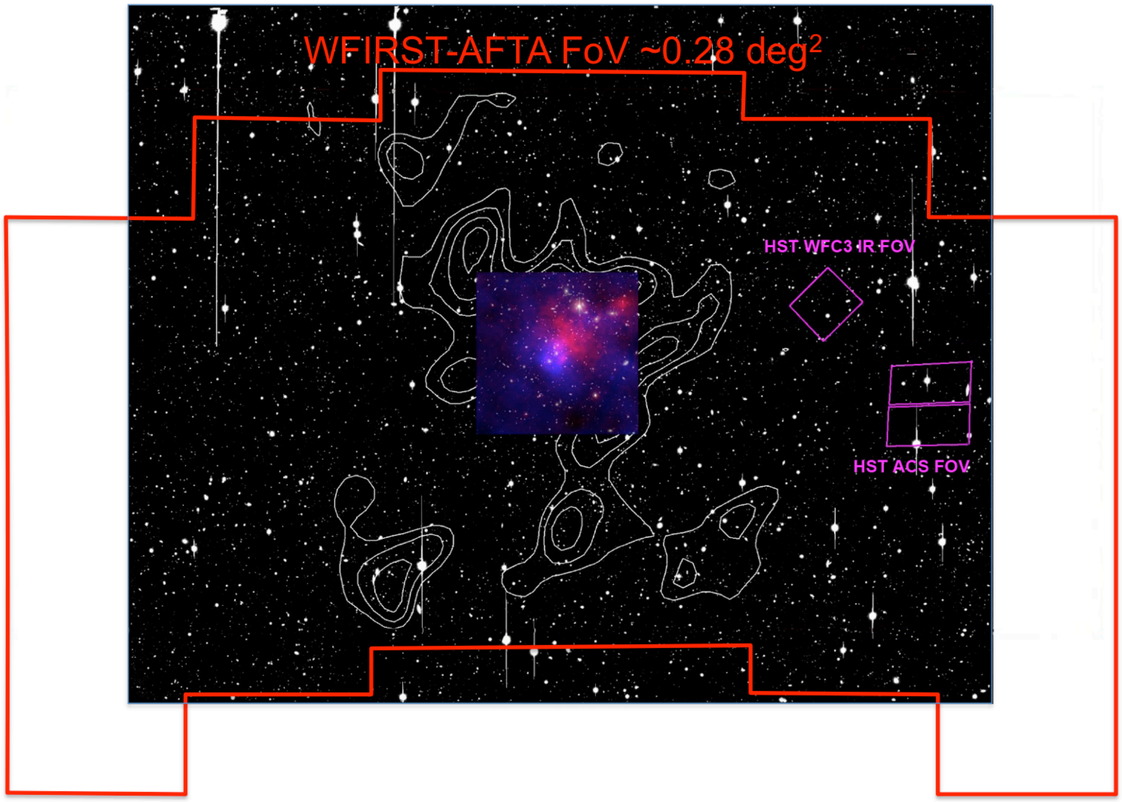


Figure 2-25: Dark matter and cluster gas distribution around the merging cluster Abell 2744. Such maps require wide-field imaging. To map the entire area above with HST WFC3 or ACS would require 100 - 200 separate pointings, due to their small field of view. The same field is mapped with a single WFIRST-AFTA pointing (red outline.

tionally. The WFIRST-AFTA HLS's unique combination of spatial resolution, infrared observations and wide area will provide mass maps that have extremely high scientific impact. Mapping the large-scale distribution of dark matter can help identify and eliminate systematic effects: unexpected spatial variations in the mass distribution can hint at exciting new physics but could also indicate observational systematics not discoverable via a power spectrum measurement like the one that will be used to constrain dark energy via the HLS.

2.3.3 *Kinematics of Stellar Streams in our Local Group of Galaxies*

The HLS will sample each patch of sky twice for each filter, with the visits distributed over the duration of the full survey. This will enable absolute proper motions to be derived for a very large number of field stars. From the ground, only bright galaxies and quasars can be used as high-precision absolute astrometric reference points. Unfortunately, their low density requires either very wide-field transformations or a bootstrapping approach that measures a target star relative to its local field population and larger field population relative to fixed reference points. WFIRST-AFTA will provide access to an enormous number of slightly resolved medium-brightness galaxies, and absolute motions can be measured directly within the field of each detector array.

This local approach to measuring absolute motions has recently been used in the optical with HST to measure the absolute motions of the globular clusters, LMC, SMC, dwarf spheroidals and even M31, in addition to individual hyper-velocity and field stars. The strategy involves constructing a template for each galaxy so that a consistent position can be measured for it in exposures taken at different epochs. This template can be convolved with the PSF to account for any variations in focus or changes of the PSF with location on the detector array. The same approach that worked with HST in the optical should work with WFIRST-AFTA in the IR. The WFIRST-AFTA detector array pixel scale and sensitivity should be very similar to that of HST's WFC3/IR detector array pixel scale. HST images of the Ultra Deep Field through the F160W (1.6 micron) pass-

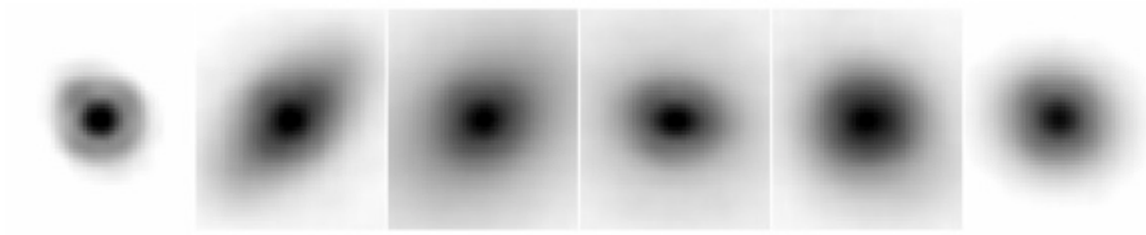


Figure 2-26: A stellar PSF (left most image) compared to 5 galaxy images from the Ultra Deep Field as seen with HST's WFC3/IR detector - an imager with similar resolution to the WFIRST-AFTA imager. Each of these objects enables a position uncertainty good to 0.02 pixel (2 mas) in a *single* exposure.

band show that there should be about 30 galaxies per square arcminute for which WFIRST-AFTA could measure a position to better than 5 mas in a 360 second exposure (see Figure 2-26). This gives us about 500 reference objects in each WFIRST-AFTA detector array, enabling us to tie down the absolute frame to better than 0.5 mas in each exposure. The HLS 2-year baseline would allow absolute motions to be derived with systematic accuracies of about 125 μas/year. Folllow-on GO programs could extend this baseline to 5-years, enabling accuracies of about 50 μas/year. There are about 10 stars in the ~5 square arc-minute FoV of the UDF that can be measured with this accuracy, implying about 18 million halo stars in the high-latitude imaging survey.

Gaia will overcome the sparse-galaxy issue by measuring positions across the entire sky in a single global solution, but it will not be nearly as sensitive as WFIRST-AFTA. Gaia will achieve the above precision for G stars down to about V = 17, which allows the plentiful turnoff-star population to be probed out to about 2.5 kpc. WFIRST-AFTA will achieve the above precision for G stars down to V=20 and will allow turnoff stars to be probed out to about 10 kpc (see Figure 2-27). There are about five times more stars within one magnitude of the turnoff than there are on the entire giant branch, and ten times more within two magnitudes of the turnoff, so WFIRST-AFTA's depth will allow us to probe well beyond the thin disk with very good statistics.

The typical dispersion of stars in the halo is ~150 km/s, but since stream stars were gently stripped from objects with low dispersion, they typically have motions that are coherent to better than 10 km/s. Such 225:1 concentrations in phase space are easy to detect. Current ground-based stream studies are focused on HB/RGB/SGB stars, but access to the much more plentiful stars below the turnoff will increase by more than an order of magnitude the number of stars we have access to and the distance out to which we can study streams. By comparing the physical locations of streams and the motions within them, we can tease out the structure of the Galactic potential.

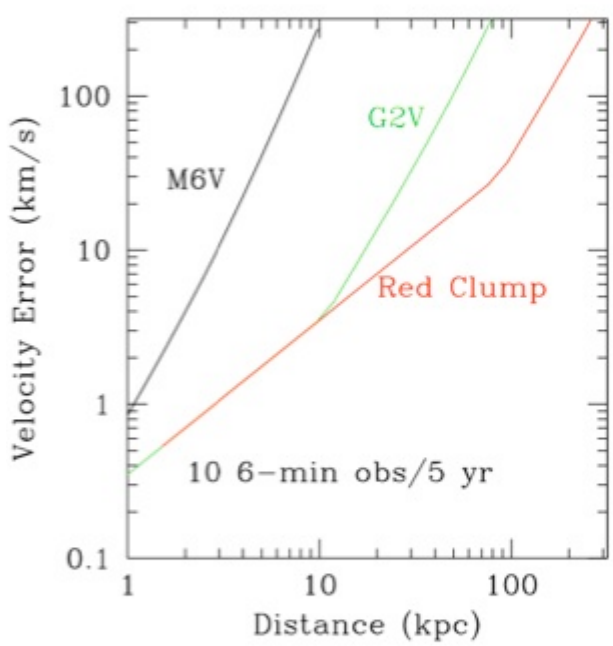


Figure 2-27: Velocity uncertainty of different components of the Galaxy's stellar population as a function of distance from the Sun. These errors are typical of what WFIRST-AFTA can achieve over a 5-year period. As part of its high latitude survey, WFIRST-AFTA will be able to measure parallax distances to over 200 million stars and will be able to track the motions of stars much fainter than those observed by the Gaia satellite.

These stellar streams are powerful probes of the distribution of dark matter in our Galaxy and can determine whether our Galaxy has the large number of million solar mass subhalos predicted in the cold dark matter model. If the dark matter is not cold, but warm as suggested in some theories, these massive subhalos will not exist, and the stellar streams will be smooth and unperturbed (Johnston et al. 2002).

2.3.4 *Discovering the Most Extreme Star Forming Galaxies and Quasars*

The supermassive black holes that sit in the centers of galaxies power active galactic nuclei (AGNs). There has been growing evidence that these supermassive black holes not only power these AGNs and quasars, but also play an essential role in the evolution of galaxies: (1) the stunning correlation between the masses of supermassive black holes, the AGN, with the masses of their host galaxies over a wide range of galaxy masses and galaxy types (Gebhardt et al. 2000, Ferrarese & Merritt 2000) and (2) numerical simulations show that star formation in the most massive galaxies must be suppressed by processes more powerful than supernovae and stellar winds in order to simulate galaxies resembling those we actually observe (e.g. Borgani et al. 2004). The most massive galaxies in the universe are the most affected by AGN feedback – they host the most massive black holes, and their star formation was truncated drastically about 10 billion years ago. The more massive the galaxy, the earlier this truncation happened, a trend called "downsizing" (Heavens et al. 2004, Thomas et al. 2005, Neistein et al. 2006). With the HLS spectroscopic survey, we will be able to probe the epoch of downsizing of the most massive galaxies in the universe: z~2, which is the peak of the star formation density and near the peak of the quasar activity.

The HLS spectroscopic survey will enable the largest census yet performed of powerful emission-line galaxies and quasars up to (and possibly beyond) lookback times of ~90% of the current age of the universe. It will enable us to better understand the relationship between the accretion of matter by active galactic nuclei and the star formation occurring in the most massive galaxies in the universe. Several key star formation indicators in the spectral range of the grism will be usable to track star formation rates (SFR) in galaxies over the range $1 < z < 4.2$. The H-alpha line will probe SFR that are at least 10 – 20 solar masses per year over the range $1 < z < 2$, a cosmic epoch that is particularly challenging to study from the ground. The estimated surface density of such galaxies is about 10^4 per square degree, implying the total survey will produce H-alpha measurements for over nearly 20 million galaxies. The OII[3727] line can be detected from star-forming galaxies lying between $z = 2.6$ and $z = 4.2$, but only for those systems with SFR in excess of ~200 solar masses per year. While such systems are less common, the WFIRST-AFTA HLS spectroscopic survey will allow us to accurately determine their space density over the key redshift range where the cosmic star formation rate has reached its highest value. At cosmic times from 200 – 600 million years since the Big Bang ($8 < z < 15$), Lyman-alpha emitting galaxies may be detectable if they have "attenuated" SFRs of at least 100 – 200 solar masses per year.

The HLS grism survey will discover ~2600 $z > 7$ quasars with H magnitude < 26.5, with an estimated 20% of those quasars being at $z > 8$. These are the luminous quasars whose existence tracks the assembly of billion solar mass black holes a mere few hundred million years after the Big Bang, and whose light can potentially illuminate the transition of intergalactic gas in the universe from neutral to an ionized phase (Gunn-Peterson effect Gunn & Peterson 1965, Becker et al. 2001, Djorgovski et al. 2001, Fan et al. 2003). Such quasars must be found before their visible red light can be studied at high spectral resolution with JWST and ground-based 10-30-m class telescopes. The fainter quasars are particularly important to find because their region of local influence is smaller (e.g. Goto et al. 2011). These are the same quasars that would be used to identify CIV in the IGM allowing the study of the chemical enrichment as a function of cosmic time. The HLS grism survey will also be a rich source of backlight targets for higher resolution spectroscopy with other facilities.

2.3.5 *HLS-Embedded Deep Fields Option*

The notional design for the HLS calls for a uniform depth survey of about 2200 deg^2 in four passbands. The survey depth and area are dictated by the requirement to ensure an ample yield of usable galaxies for the weak lensing component of the dark energy survey. However, there are advantages to consider a version of the HLS that, in addition to a wide area survey of at least 2000 deg^2, includes one or more regions of significantly greater depth (1.5 to 2.5 magnitudes fainter) but covering more modest areas (e.g., 20 to 50 deg^2). At depths of +1.5 mag fainter than the nominal HLS wide field survey, a WFIRST deep field, performed as part of a larger HLS, would provide NIR depths that are comparable to the *ugriz* limits of ~28 AB mag expected for the LSST survey. This deeper imaging would provide a superb training set for assessing the reliability of the photometry and shapes of galaxies near the limits of the wide-field shallower HLS and would also improve the robustness of such measurements for the full LSST survey. These HLS-embedded deep fields could cover 4 to 10 times more sky area than the WFIRST SN deep fields and in more passbands. An extra 1.5 mag of depth would require exposure times that are ~10x longer than the nominal HLS exposures. A deep tier to the HLS of 20 to 50 deg^2 could thus be completed in about 10% – 20% of the time required for the full HLS. The exciting science enabled by including selectively targeted deep fields within the HLS includes (1) Measurements of the large-scale distribution and clustering of $z > 7$ galaxies and of high-redshift AGN, (2) Characterizing the properties of $z > 2$ galaxy clusters, (3) Ultra-faint surveys of the Milky Way Halo, and (4) Tracking AGN and QSO variability.

2.5 Galactic Bulge Field: General Astrophysics

The WFIRST-AFTA microlensing survey toward the Galactic bulge will potentially enable a very broad range of astrophysics, well beyond the primary goal of completing the census of cold exoplanetary systems. This enormous additional potential of the survey follows from the large number of stars being surveyed, the large number of measurements that will be made for each of these stars, and the quality of the photometric and astrometric precision that will be achieved for these stars, which will range from good to exquisite, depending on the brightness of the star and the level to which systematics can be controlled.

In this section, we provide a sampling of the range of the science that may be possible, including auxiliary science from the microlensing events themselves, stellar astrophysics, Galactic structure, and solar system science. In §2.4.2.9 we discussed the additional exoplanet science that can be obtained from the survey from the detection of tens of thousands of transiting giant planets.

There are a few things to note. First, this overview contains an admixture of potential applications, some of which have been worked out in some detail, and some of which have merely been suggested, with no real effort made to demonstrate their viability. These latter topics provide an important line of future work to vet their applicability. Second, we are primarily concentrating on auxiliary science that can potentially be extracted without changing the microlensing survey specifications. However, some of these applications will require new data reduction algorithms or precursor or follow-up observations in order to extract their full science potential. Finally, we note that several of these applications require excellent-to-exquisite control of systematics in the photometric and astrometric measurements. The extent to which these systematics can be controlled at the levels required is unclear. However, we note that it is likely that the microlensing survey itself will provide the best opportunity to characterize and remove these systematics by its very nature of requiring many repeated observations of a large number of point sources. We discuss this in a bit more detail below.

2.5.1 *Properties of the Survey*

We begin by summarizing the properties of the microlensing survey. As currently envisioned, the survey will consist of ten contiguous fields (see Figure 2-29) between Galactic longitudes of roughly -0.5 and 1.8, and Galactic latitudes of -1 and -2.2, for a total area of roughly 2.81 square degrees. The fields will be imaged in six 72-day campaigns spread over the nominal six-year mission lifetime, but due to losses from the moon moving through our field, the total survey time will be 357 days. During each campaign, each field will be observed every 15 minutes in a wide (W149) filter that spans 0.9-2 microns, and once every 12 hours in a bluer filter (Z087) that spans 0.77-0.97 microns. The exposure time per epoch in these filters with be 52 secs for W149 and 290 secs for Z087. Because of the layout of the detectors in the focal plane, ~85% of the survey footprint is observed for the full time while the remaining 15% is observed only in either the Spring or Fall seasons. For stars in the primary area, there will be a total of ~33,000 epochs in W149 and ~700 in Z087. The single-measurement photon-noise photometric and astrometric precisions will be ~0.06% and 0.05 mas for the one million stars with H_{AB}<14.0, ~0.4% and 0.7 mas for the 12 million stars with H_{AB}<19.6, and ~1.2% and 1.7 mas for the 56 million stars with H_{AB} < 21.6. WFIRST-AFTA will collect ~2.5 billion photons for a star with H_{AB}=19.6 in the wide filter over the course of the mission.

2.5.2 *Microlensing Auxiliary Science*

The microlensing survey will be sensitive to microlensing due to isolated compact objects with masses from that of Mars to roughly 30 times the mass of the Sun. We expect 37,000 microlensing events due to known populations alone (stars and stellar remnants) (Penny et al. 2015a). Thus it will be able to measure or constrain the mass function of compact objects over roughly 8 orders of magnitude in mass. This includes free-floating planets, brown dwarfs, stars, and stellar remnants. For the upper end of this range, it will be possible to measure the masses of the objects directly from higher-order microlensing effects from the survey data itself (Gould & Yee 2014), whereas for lower-mass objects, either additional measurements may be required, or the mass function will only be determined statistically. It will also be possible to probe the binarity of these compact objects, including the companion frequency and mass ratio and separation distribution, for binaries with separations within roughly a decade of the Einstein ring radius of the primary. It will also be possible to detect tertiaries in some cases.

With the large number of microlensing events, it will be possible to measure the microlensing optical depth and event rate over the ~3 square degrees of the target fields. The optical depth, in particular, is proportional to the integral of the total mass density of compact objects along the line of sight, and as such is a very sensitive probe of the mass distribution of the Galactic disk and the Galactic bulge (Griest et al. 1991, Kiraga & Paczynski 1994). In particular, the optical depth can be used to probe the total mass, shape, and orientation of the Galactic bar (Paczynski et al. 1994, Zhao et al. 1995). The event rate is additionally sensitive to the velocity distribution and mass function of the compact objects in the stellar populations.

Roughly 5% of all microlensing events exhibit clear signatures of caustic crossings due to the lens being composed of a nearly equal mass binary with a projected separation near the Einstein ring radius (Alcock et al. 2000). With the near-continuous sampling and 15 minute cadence, nearly all of the caustic crossings of giant sources (which last ~10 hours) will be well-resolved, thus allowing one to measure the limb-darkening for these stars in the wide filter (Albrow et al. 1999).

2.5.3 *Stellar Astrophysics*

With a typical total exposure time of ~20 days per field in the wide (W149) filter and ~2.3 days per field in the blue (Z087), it will be possible to construct a very deep color-magnitude diagram over the ~3 square degrees targeted by the microlensing survey. The limiting magnitude in both filters is uncertain, but will likely be set by confusion, rather than photon noise. Nevertheless, it should be possible to measure the luminosity function of stars in the bulge down to nearly the bottom of the main sequence, as well as identify unusual stellar populations down to quite faint magnitudes, including potentially young stars, blue stragglers, stellar clusters, cataclysmic variables, and white dwarfs. Such studies will be enhanced by the fact that it will be possible to measure the proper motion and potentially parallax for a significant subset of the sources in the bulge field, thus allowing for discrimination between bulge and disk populations. Such science would be significantly enhanced by observations in one or more bluer filters at similar angular resolution, either with WFIRST-AFTA itself, or with Hubble Space Telescope precursor observations.

Of course, it will be possible to identify many variable sources in the fields, over a broad range of amplitudes and time scales. Variables can be identified with amplitudes greater than a few millimagnitudes, with time scales from tens of minutes up to the ~6-year duration of the mission. This will allow for the identification of a large number of stellar flares, eruptive variables, and pulsating variables, thereby enabling a wide variety of stellar astrophysics.

Perhaps most exciting, however, is the potential for WFIRST-AFTA to do asteroseismology on, and thus determine the mass and radii of, a significant number of stars in the survey area. Gould et al. (2015) estimate that WFIRST-AFTA will obtain asteroseismic information on the roughly one million giant stars with $H_{AB} < 14$ in the target fields (see Figure 2-51). They also estimate that WFIRST-AFTA will obtain precise (~0.3%) distances to these stars via parallax. This is the only way to measure masses and radii of a large number of bulge giants, which are postulated to have abundances that are substantially different than local populations.

2.5.4 *Galactic Structure*

WFIRST-AFTA will obtain photon-noise limited parallaxes of <10% and proper motion measurements of <0.3% (0.01 mas/yr) for the ~56 million bulge and disk stars with H_{AB}<21.6 in its field of view. When com-

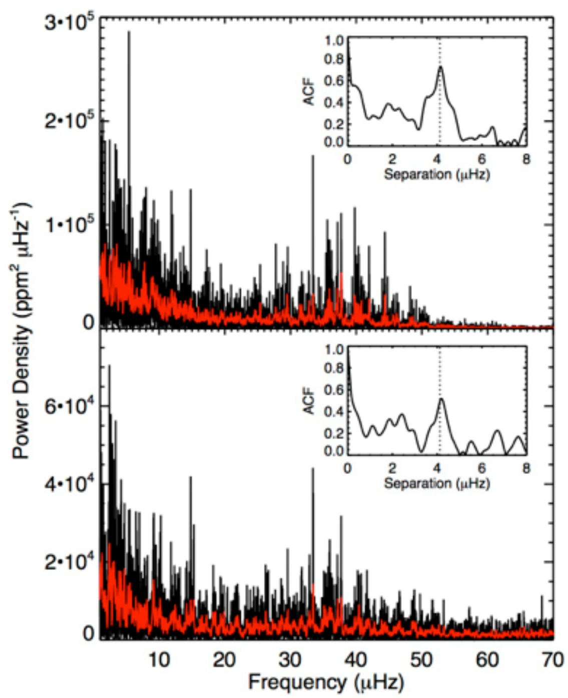


Figure 2-51: Asteroseimologic spectrum for KIC2836038 as observed by Kepler (top panel) and as simulated for WFIRST-AFTA (bottom panel). Red lines are the power spectra smoothed with a boxcar width of 1μHz. The insets show the autocorrelation function of the smoothed power spectrum between 20 − 60μHz after correcting the granulation background. The dotted line marks the published value for the large frequency separation Δυ. WFIRST-AFTA will be able to measure Δυ for the roughly one million giant stars with H_{AB}<14 in the microlensing survey fields. From Gould et al. (2015).

bined with multicolor photometry in at least one bluer filter (to complement the measurements in W149 and Z087), it will be possible to estimate the effective temperature, metallicity age, luminosity, and foreground extinction for all of these stars. This is a larger number of stars with such measurements than will be achievable with Gaia, although of course the population of stars will be very different and complementary to that obtained by Gaia. From this dataset, it will be possible to extract an enormous amount of Galactic structure science, including a determination of the bulge mass and velocity distribution (including bar structure), the stellar density and velocity distribution of the Galactic disk, the metallicity and age distribution of the disk and bulge, and the three-dimensional distribution of dust along the line of sight toward the bulge fields.

2.5.5 *Solar System Science*

Gould (2014b) estimates that WFIRST-AFTA will detect ~5000 Trans-Neptunian objects (TNOs) down to absolute magnitudes of H_{AB}~29.6 (corresponding to diameters of D~10 km) over ~17 square degrees, enabling a precise determination of the size distribution of TNOs down to, and substantially below, the collisional break at ~100 km (see Figure 2-52). Furthermore, the orbital elements of these objects will be measured to fractional precisions of a few percent, allowing for dynamical classification into the canonical classical, resonant, and scattered populations, as well as identification of objects with new or unusual orbits. Binary companions to these objects will be discovered down to fainter magnitudes of H_{AB}~30.4, corresponding to diameters of 7 km, thus allowing for the determination of the binary size and separation distribution as a function of dynamical class. Finally, there will be ~1000 occultations of stars due to these TNOs during the bulge survey (Figure 2-52), allowing for a statistical estimate of the size and albedo distributions of TNOs as a function of magnitude and dynamical class. The several dozen TNOs discovered with H_{AB}<25 (D>100 km) will occult multiple stars during the survey, allowing for crude constraints on the shapes of these large TNOs.

2.5.6 *Related Projects and Precursor Observations*

The Galactic bulge survey of WFIRST-AFTA is unique in terms of its resolution, area, exposure time, total number of observations, and total number of sources. Therefore, it is essentially guaranteed to produce revolutionary science in all of the areas listed above. Nevertheless, there are a number of existing, planned, or proposed surveys with similar science

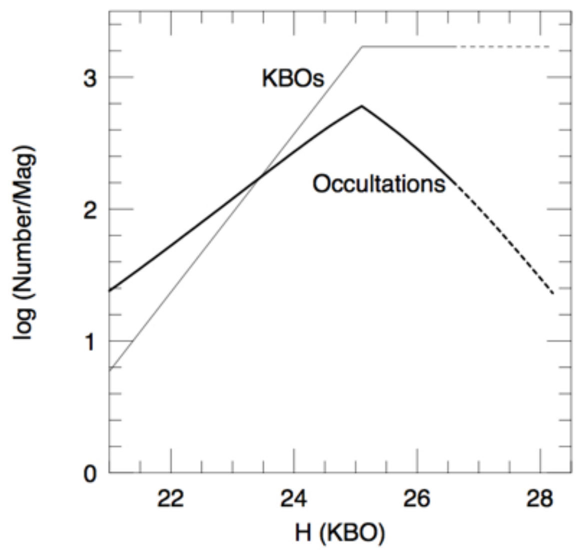


Figure 2-52: The expected number of Trans-Neptunian Object (or KBO) detections (light solid line) and TNO occultations (heavy solid line) per magnitude that can be detected from the WFIRST-AFTA microlensing survey. Magnitudes here are in the Vega system, which is related to the AB system used in the text by $H_{Vega} = H_{AB} – 1.39$. WFIRST-AFTA will detect TNOs significantly fainter the current survey limits of H_{Vega} = 26. KBOs with H_{Vega} < 23.5 will occult more than one star during the survey, potentially allowing for crude constraints on their shape. From Gould (2014b).

goals, which will motivate, enhance, and/or complement the science potential of the WFIRST-AFTA bulge survey. These include the ground-based near-infrared: Vista Variables in the Via Lactea (VVV) Survey (Minniti et al. 2010), and optical: Blanco DECam Bulge Survey (BDBS), the Hubble Space Telescope Galactic Bulge Treasury Program (Brown et al. 2009) and Sagittarius Window Eclipsing Extrasolar Planet Search (SWEEPS) (Sahu et al. 2006), and the proposed Japan Astrometry Satellite Mission for Infrared Exploration (JASMINE) (Gouda et al. 2005). The science goals of the JASMINE mission concepts, in particular, have significant overlap with the auxiliary science topics listed above[9].

As discussed in Appendix G and advocated in the recent NASA ExoPlanet Analysis Group (ExoPAG) Study Analysis Group 11 report, precursor observations with HST of some or all of the area targeted by the WFIRST-AFTA bulge survey would significantly enhance many of the auxiliary science goals discussed above. HST imaging is well matched to that of WFIRST-AFTA in terms of resolution and sampling,

[9] http://www.scholarpedia.org/article/JASMINE

while providing both bluer filters and the possibility of obtaining a longer time baseline of observations than is possible with WFIRST-AFTA alone. Precursor observations with HST could then be combined with the WFIRST-AFTA data to provide color-magnitude diagrams of the survey area in several colors, confirmation of low signal-to-noise ratio astrometric and proper motion measurements made internal to the WFIRST-AFTA bulge survey, and to provide locations and colors of all the point sources with higher resolution and fidelity than is possible with WFIRST-AFTA alone.

2.5.7 *Detector Characterization*

With its ~33,000 dithered images of ~10^8 point sources with known colors and nearly constant fluxes, the bulge survey will likely provide one of the best datasets with which to characterize WFIRST-AFTA's wide-field imager. In particular, we imagine that the bulge survey pipeline will measure the fluxes and positions of all of the stars in the field of view, while simultaneously measuring the response function of each of the WFI pixels, and how this response function varies with time. This time-variable response function will then be used to calibrate and remove detector artifacts such as persistence, inter-pixel capacitance, reciprocity failure, non-linearity, and intrapixel response variations.

Jason Kalirai (STScI), jkalirai@stsci.edu

The Infrared Color-Magnitude Relation

Background

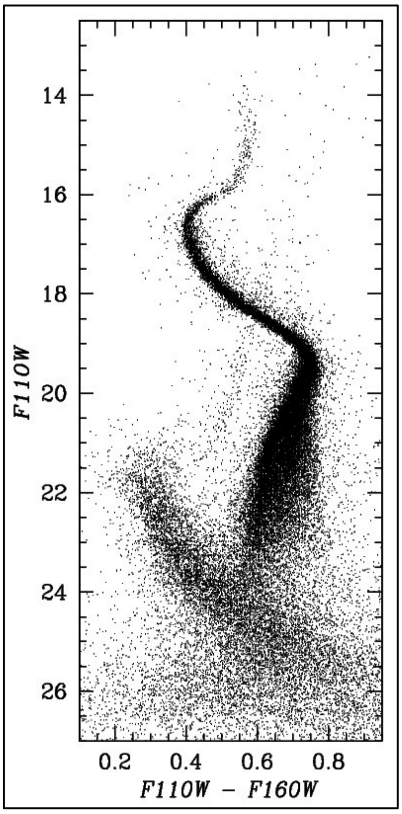
Star clusters in the Milky Way (MW) have served as the primary tools to measure the color-magnitude relation of stars, and to calibrate its dependency on stellar properties such as age and metallicity. This relation is a key input to test stellar evolution models, and in turn to carry out population synthesis studies that aim to interpret the integrated light of astrophysical sources across the Universe (e.g., Bruzual & Charlot 2003). For decades, this work has primarily focused on the interpretation of visible-light color-magnitude diagrams (CMDs).

WFIRST

WFIRST will enable high-precision IR CMDs of stellar populations. The figure below illustrates the morphology of the IR CMD of the globular cluster 47 Tuc, from a 3 orbit (depth) exposure with the WFC3/IR camera on HST (Kalirai et al. 2012). The sharp "kink" on the lower main sequence is caused by collisionally induced absorption of H_2. Unlike the visible CMD, the inversion of the sequence below the kink is orthogonal to the effects of distance and reddening, and therefore degeneracies in fitting fundamental properties for the population are largely lifted. The location of the kink on the CMD is also not age-sensitive, and therefore can be used to efficiently flag 0.5 Msun dwarfs along any Galactic sightline with low extinction. A WFIRST two-stage survey will first establish the IR color-magnitude relation and the dependency of the "kink" on metallicity through high-resolution, deep imaging of Galactic star clusters. Second, this relation can be applied to field studies to characterize the stellar mass function along different sightlines, the dependency of the mass function on environment, and to push to near the hydrogen burning limit in stellar populations out to 10's of kpc.

Key Requirements

Depth – Well dithered exposures extending down to the H burning limit in clusters with [Fe/H] = -2.2 to 0.0 (i.e., 10 kpc)
Field of View #1 – Single pointings for globular clusters covering appreciable spatial extent
Field of View #2 – Wide field survey of Galactic plane sampling over star forming regions and spiral arms
Cadence – One image per galaxy
Wavelength Coverage – Two NIR filters for color-magnitude diagram analysis



Caption: IR color-magnitude diagram for the nearby globular cluster 47 Tuc, constructed from a 3 orbit (depth) observation with HST/WFC3/IR (Kalirai et al. 2012). The kink in the lower main-sequence of the cluster is caused by H_2 opacity. The fainter main-sequence represents stars from the background SMC galaxy.

Kailash Sahu (STScI), ksahu@stsci.edu

Detecting Neutron Stars and Black Holes through Astrometric Microlensing

Stars with initial masses of ~8–20 M⊙ are expected to end their lives as NSs, and those with greater masses as BHs (Heger et al. 2003). As much as 10% of the total mass of our Galaxy should be in the form of NSs and BHs (Oslowski et al. 2008). However, a vast majority of stellar remnants are expected to be single, either primordially or due to disruption of binaries by supernova (SN) explosions (Agol & Kamionkowski 2002). Such isolated massive remnants are extremely difficult to detect directly, and in fact no isolated BH has ever been unambiguously found within our Galaxy. Mass measurements of NSs and BHs mostly come from observations of binary systems which show NS masses to be concentrated around 1.4M⊙, and BH masses to be a narrow Gaussian at 7.8 ± 1.2 M⊙ (Ozel et al. 2010). However, theoretical models (Fryer & Kalogera 2001) predict that the compact-remnant distribution should be a continuous distribution from 1.2 to 15M⊙. The discrepancy between the observed and predicted mass distribution of BHs is generally attributed to the fact that, BHs in binaries are a biased and minority sample. What is missing is an unbiased mass distribution for isolated stellar remnants.

Astrometric Microlensing: Isolated stellar remnants can be detected through microlensing. The microlensing survey programs such as OGLE and MOA have so far detected more than 8000 microlensing events. If stellar remnants constitute a few percent of the total mass, many of these observed microlensing events must be due to stellar remnants. However, microlensing light curves are degenerate with respect to the mass, velocity and distance of the lens. A route to resolving these degeneracies arises from the fact that microlensing, in addition to amplifying the brightness of the source, produces a small shift in its position (Dominik and Sahu, 2000). Thus if high-precision astrometry is added to the photometry, the deflection of the source image can be measured, and thus the mass of the lens determined unambiguously. The sizes of the astrometric shifts, however, are such that they require a 2m-class space based telescope.

WFIRST/NRO: Astrometric precisions of order 300 microarcsec have been achieved with HST/WFC3, so we expect that a similar astrometric precision is achievable with NRO. The maximum astrometric deflection caused by a 0.5 M⊙ lens at 2 kpc is 400 μas (see Figure 1). This implies that NRO can measure the astrometric deflection caused by almost all NSs and BHs.

In order to unambiguously measure the mass of the lens, the remaining quantity to be determined is the distance to the lens. Fortunately, the timescales of microlensing events caused by NSs and BHs are long (>30 days), so photometric observations of these events with NRO will provide clear measurements of the parallax distances.

From our HST observations towards the galactic bulge, we estimate that NRO will enable monitoring of about 20 million stars towards the Galactic bulge in a single pointing, covering about 0.25 square degrees. By monitoring 10 fields in the Galactic bulge, 200 million stars can be easily monitored continuously. Since the optical depth towards the Galactic bulge is about 3 * 10(-6), this will lead to the detection of 600 microlensing events at any given time, and 5000 events per year. Taking the current statistics of microlensing events, about 10% of them (500 events) are expected to have $T_E > 30$ days, and at least three dozen of them are expected to be due to NSs and BHs. Observations taken over 5 years will lead to the detection and mass measurement of well over 100 NSs and BHs. These measurements will provide (i) constraints on SN/GRB explosion mechanisms that produce NSs and BHs, and (ii) a quantitative estimate of the mass content in the form of stellar remnants.

Scott Gaudi (OSU), gaudi@astronomy.ohio-state.edu, Matthew Penny (OSU), Naoteru Gouda (NAOJ), Yoshiyuki Yamada (Kyoto U.)

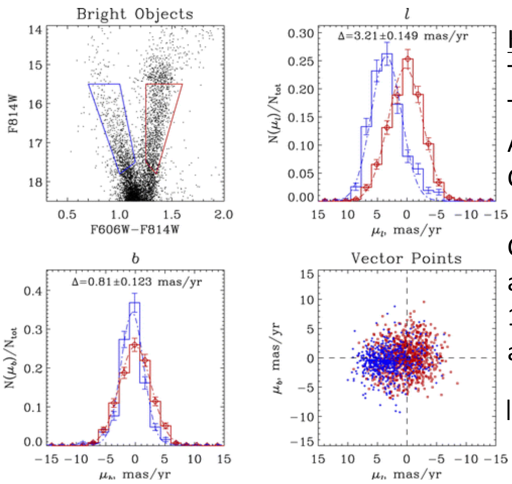
Proper Motions and Parallaxes of Disk and Bulge Stars

Background

Measurements of the kinematics and three-dimensional structure of stars in the Galactic Bulge and inner disk allow for the determination of the dynamical mass in these populations, and provide important clues to their formation and evolutionary history. The dominant formation mechanism of bulges in the universe (i.e., secularly-grown pseudobulges versus merger-driven classical bulges) remains poorly understood (i.e., Kormendy & Kennicutt 2004), and our bulge provides the nearest and thus most accessible example with which to test the predictions of various formation models. Furthermore, the above measurements can provide useful information with which to clarify the evolutionary processes of the supermassive black hole at the Galactic center and its relation to the evolution of the Galactic bulge. Nevertheless, there are a number of challenges to obtaining such measurements for the bulge, in particular the small proper motions and parallaxes, and the large and variable extinction towards the bulge. To date, information on the kinematics of bulge stars has been limited to primarily radial velocities, luminous stars or stars with large proper motions, or a few deep but narrow pencil-beam surveys with HST. Direct geometrical distances to individual stars in the bulge have been essentially unavailable.

WFIRST-AFTA

WFIRST-AFTA will enable high-precision proper motion measurements and parallaxes to essentially all the ~56 million bulge and foreground disk stars in the ~2.8 square degree microlensing survey field-of-view with magnitudes of $H_{AB}<21.6$. For the exoplanet survey, WFIRST-AFTA will achieve photometric precisions of ~1.2% per 52s observation for stars with $H_{AB}<21.6$. Assuming a resulting astrometric precision of a σ_{AST} ~ 2 mas per observation (i.e., σ_{AST} ~ FWHM/SNR~0.14”X0.012~1.7 mas), and N~33,000 observations, the final mission uncertainty on the measured proper motions over a T~5 year baseline will be $\sigma_{\mu\mu}$ ~ $(12/N)^{1/2}(\sigma_{AST}/T)$ ~ 0.01 mas/yr. The typical proper motion of a star in the bulge is μ~100 km/s/(8000 pc) ~ 3 mas/year, and thus individual stellar proper motions will be measured to ~0.3%. Similarly, the fractional uncertainty on the parallax of a star in the bulge will be σ_Π/Π ~ $(\sigma_{AST}/\Pi)(2/N)^{1/2}$ ~ 10%, where Π ~ 1/8 mas is the typical parallax of a star in the bulge. Note that these estimates assume that systematic uncertainties can be controlled to better than ~0.01 mas, or 1/1000th of a pixel. Because these observations will be taken in the NIR, they will reach below the bulge MS turnoff and will be relatively unaffected by extinction. The occasional observations in bluer filters will help distinguish between bulge and disk populations. With these measurements, AFTA-WFIRST will provide unprecedented measurements of the kinematics and structure of the bulge.



Key Requirements

Total Span of Observations: T~ 5 years

Total number of epochs: ~33,000

Astrometric precision per epoch for $H_{AB}<21.6$ of ~2 mas

Control of systematic errors to better than ~0.01 mas

Caption: CMD and bulge, disk proper motion distributions from an HST study using ASC/WFC (Clarkson et al 2008, ApJ, 684, 1110). WFIRST-AFTA will provide individual proper motions that are roughly an order of magnitude more accurate than HST.

Gurtina Besla (Columbia) gbesla@astro.columbia.edu; Roeland P. van der Marel (STScI) marel@stsci.edu

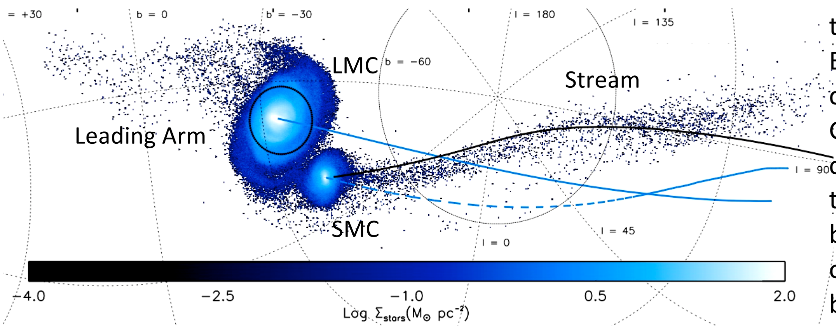
The Detection of the Elusive Stellar Counterpart of the Magellanic Stream

Background

As the most massive satellites of the Milky Way (MW), the Large and Small Magellanic Clouds (LMC and SMC) play an important role in our still developing picture of the buildup and evolution of the Local Group. In the ΛCDM paradigm, MW-type halos are expected to have built up the majority of their mass by the accretion of LMC-type subhalos; *the evolution and disruption of the Magellanic System is thus directly relevant to our understanding of how baryons are supplied to the MW.*

The Magellanic Clouds (MCs) are undergoing substantial gas loss, as is evident by the stream of H I trailing (the Magellanic Stream) and leading the MCs (the Leading Arm), in total stretching 200° across the sky (Mathewson et al. 1974; Nidever et al. 2010). Despite extensive modeling and multi-wavelength studies of the system, the dominant mechanism for the formation of this extended gas distribution is unknown. Leading theories are: tidal stripping of the SMC (by the MW, Gardiner & Noguchi 1996; or by the LMC alone, Besla, Hernquist & Loeb 2012; see Figure) or hydrodynamic processes (ram pressure stripping, Mastropietro et al. 2005; stellar outflows, Nidever et al. 2008). *Distinguishing between these formation scenarios is critical to the development of an accurate model for the orbital and interaction histories of the MCs with each other and with the MW.*

In the tidal models, stars are removed in addition to gas. In contrast, stars are not removed in any of the hydrodynamic models. *The detection of stellar debris in the Magellanic Stream would prove conclusively that the Stream is in fact a tidal feature, ruling out models based on hydrodynamic processes.* Although previous searches for stars in sections of the Stream have yielded null results (e.g., Guhathakurta & Reitzel 1998, Bruck & Hawkins 1983), the predicted optical emission from the stellar debris in e.g., the Besla, Hernquist & Loeb 2012 model (>32 mag/arcsec^2; V-band), is well below the observational limits of these prior surveys.



Caption: The stellar counterpart to the Magellanic Stream from Besla, Hernquist & Loeb (2012); color indicates stellar density. Galactic coordinates are overplotted. The past orbits of the MCs are indicated by the blue lines. The current location of the gaseous Stream is marked by the solid black line. Densities < 0.01 M⊙/pc^2 are typical.

WFIRST

WFIRST will enable a uniform, wide-field, deep survey across sections of the gaseous Magellanic Stream to ascertain the existence of a faint stellar counterpart. Using MC main sequence stars as tracers of stellar debris has been proven to reach equivalent population surface brightness of >33 mag arcsec^{-2} based on CTIO Mosaic-2 imaging (Saha et al. 2010); with WFIRST's wider field of view, larger surface brightness limits can be reached. Furthermore, a wide area survey will uncover any spatial offsets between the gaseous and stellar stream, which may occur owing to hydrodynamic gas drag.

Key Requirements

Depth – NIR mag limit of >23.5 (to reach I-band surface brightness >33 mag/arcsec; cf. Saha et al. 2010)
Field of View – contiguous mapping across the Stream (~ 20 degrees)
Wavelength Coverage – Two NIR filters for color-magnitude diagram analysis

Christopher J. Conselice (Nottingham), conselice@nottingham.ac.uk

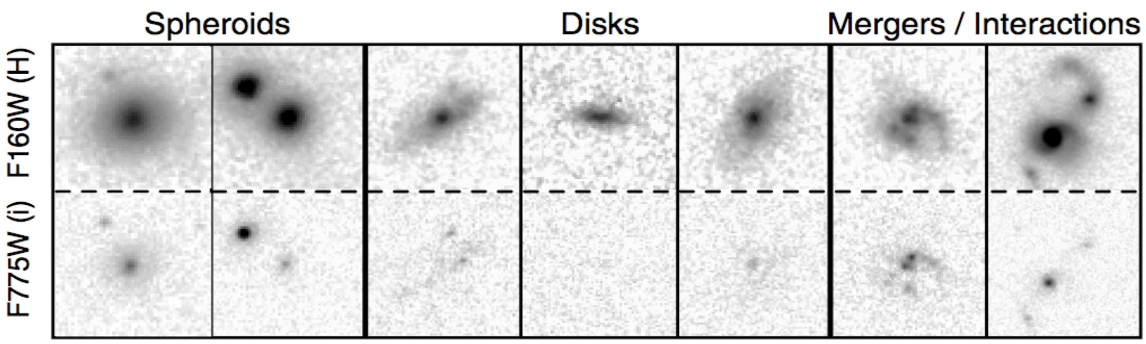
Galaxy Structure and Morphology

Introduction

There are several major issues within galaxy formation and evolution where WFIRST/NRO telescope will make a big impact. One of these critical aspects of a WFIRST/NRO mission will be to study galaxy structure - i.e., sizes and morphologies of galaxies on a larger area than is possible with Hubble Space Telescope, or from the ground using adaptive optics. Galaxy structure is critical for deciphering the processes which drive galaxy assembly and allow us to move beyond simply counting properties of galaxies to study directly their assembly processes through e.g., mergers, gas accretion, and star formation. Yet this has proved difficult to do for galaxies at z > 1. Properties such as the asymmetry of a galaxy's light, its concentration, and the clumpy nature of this light all reveal important clues to the galaxy formation process.

WFIRST/NRO

The Hubble Space Telescope has shown how powerful this resolved structural approach is for understanding galaxies. Only very small fields have been observed with Hubble, and most of those within the observed optical. To study the formation of galaxies will require high resolution imaging in the near-infrared, which WFIRST/NRO will provide, especially if a large 2.4m mirror is utilized. This will bring the pixel size down to 0.13-0.11 arcsec, allowing us to perform WFC3 type surveys over at least a few degrees, which will be much larger than any near infrared survey of this type for distant galaxies yet performed. This will allow us to directly measure in an empirical way how galaxies, and the stars within them, assembled over most of cosmic history to within a Billion years of the big bang.



Caption: How infrared light (F160W) shows features of distant galaxies not seen in optical light (F775W) (from Kocevski et al. 2011)

Key Requirements

Depth – Deep enough to examine galaxy structure down to ⅕ L* up to z = 4

FoV – Large survey areas to obtain statistically representative galaxy populations

Wavelength – Need > 1.6 microns to examine the rest-frame optical structures of galaxies

Daniel Stern (JPL), daniel.k.stern@jpl.nasa.gov

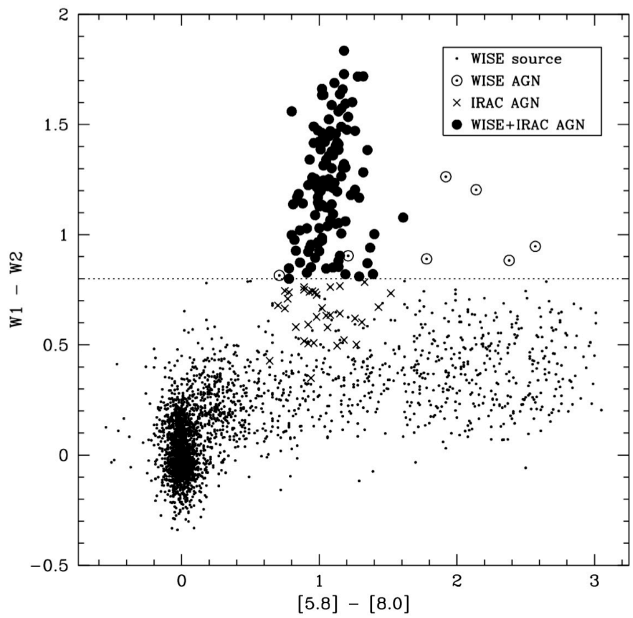
Obscured Quasars

Background

Obscured, or type-2 quasars are expected to outnumber unobscured, type-1 quasars by factors of 2-3. They are predicted by models of active galactic nuclei (AGN), and are required to explain the hard spectrum of the cosmic X-ray background. Until recently, however, our census of this dominant population of AGN has been lacking since such systems are difficult to identify in optical and low-energy X-ray surveys. High-energy missions to date have had limited sensitivity, while the recently launched NuSTAR mission has a very limited field-of-view. Mid-infrared surveys, initially with Spitzer (e.g., Lacy et al. 2004; Stern et al. 2005) and more recently with WISE (Stern et al. 2012; Assef et al. 2013) have dramatically changed the situation. In its shallowest fields, WISE identifies 60 AGN candidates per deg2 with a reliability in excess of 95%. In higher latitude, deeper parts of the WISE survey, this surface density rises to >100 AGN per deg2. Comparably powerful AGN identified in optical surveys have surface densities of only ~20 per deg2 (e.g., Richards et al. 2006), implying that WISE has finally realized the efficient identification of the dominant luminous AGN population across the full sky.

WFIRST

While mid-infrared observations with WISE have identified this dominant population of obscured, luminous quasars, WFIRST will be required to characterize its properties. Working with deep ground-based optical data, WFIRST will provide photometric redshifts for this population, something not possible from the mid-infrared data alone since the identification relies on the power-law mid-infrared spectra of luminous AGN. Photometric redshifts will allow us to probe the cosmic history of obscured black hole growth, and relate it to galaxy formation and evolution. AGN feedback is expected to play an important role shaping the present-day appearances of galaxies. Measuring the clustering amplitude of obscured and unobscured quasars will probe AGN unification scenarios. For the classic orientation-driven torus model, clustering should be the same for both populations. On the other hand, if obscured AGN are more common in merging systems, with the obscuration caused by galactic-scale material, then the clustering amplitudes are expected to differ.



Caption: Mid-infrared color-color diagram, illustrating that a simple WISE color cut of W1-W2 ≥0.8 identifies 62 AGN per deg2 with 95% reliability assuming the complete reliability of the Stern et al. (2005) Spitzer AGN selection criteria. From Stern et al. (2012).

Key Requirements

Depth – Provide robust photometric redshifts for sub-L* populations to z~2
Field of View – Wide-area surveys required for clustering analysis
Wavelength Coverage – >2 NIR filters for robust photometric redshifts

Xiaohui Fan (University of Arizona), fan@as.arizona.edu

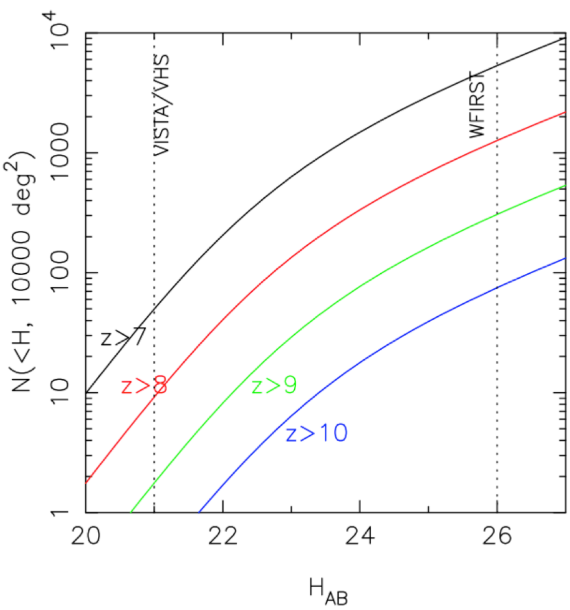
High-Redshift Quasars and Reionization

Background

Luminous quasars at high redshift provide direct probes of the evolution of supermassive black holes (BHs) and the intergalactic medium (IGM) at early cosmic time. The detection of z>7 quasars (e.g., Mortlock et al. 2011) indicates the existence of billion solar mass BHs merely a few hundred million years after the Big Bang, and provides the strongest constraints on the early growth of supermassive BHs and their environments. Spectroscopy of the highest redshift quasars reveals complete Gunn-Peterson (1965) absorption, indicating a rapid increase in the IGM neutral fraction and an end of the reionization epoch at z=6-7 (Fan et al. 2006). Current observations suggest a peak of reionization activity and emergence of the earliest galaxies and AGNs at 7<z<15, highlighting the need to expand quasar research to higher redshift.

WFIRST

While ground-based surveys such as LSST and VISTA will make progress in the z=7-8 regime in the coming decade, strong near-IR background from the ground will limit observations to the most luminous objects and to z<8. Wide-field, deep near-IR survey data offered by WFIRST will fundamentally change the landscape of early Universe investigations. Figure 1 shows the predicted number of high-redshift quasars in a WFIRST survey based on current measurements at z~6 (Jiang et al. 2008; Willott et al. 2010) and extrapolation to higher redshift with a declining number density following the trend seen at z=3-6. WFIRST should allow robust identifications of a large sample of reionization-epoch quasars up to z>10, if they exist at those epochs. WFIRST grism will provide direct spectroscopic confirmation and characterization of these high-redshift quasars, while JWST and next-generation extremely large telescope high-resolution spectroscopic observations will measure IGM and BH properties. Key questions to be addressed by these observations are: (1) when did the first generation of supermassive BHs emerge in the Universe; (2) how and when did the IGM become mostly neutral; (3) did quasars and AGNs play a significant role in the reionization process?



Key Requirements

Depth – To study sub-L* QSO populations at high redshift

Morphology – To separate stars from galaxies; one of the main contaminations is expected to be low-redshift red galaxies

Grism – For spectroscopic confirmation and characterization

Field of View – Wide-area surveys required to identify rare populations

Wavelength Coverage – >2 NIR filters to robustly identify QSOs at the highest redshifts

SECTION 3: WFIRST OBSERVATIONS, DATA PROCESSING, AND DATA PRODUCTS

Preface

As described in the SDT 2015 Report, WFIRST will make four major surveys including:

- The High Latitude Imaging Survey
- The High Latitude Spectroscopic Survey
- The Supernova Survey
- The Microlensing Survey

For each of these surveys, we give a preliminary description of the observations, required processing of the observational data, and the envisioned data products. The description of the observations given here is a lightly edited version of text in the 2015 Report, and the figures are screen captures of the figures in the 2015 Report, so the figure captions retain the same numbering system as the 2015 report.

The data description tables, data-processing, and data products from each survey have no counterpart in the 2015 Report. We have derived the needed data processing and data products based to the best of our knowledge on the scientific goals of WFIRST, expected WFIRST observational data, and on our previous experience with NASA's astrophysics space telescopes. As WFIRST is in the early stage of development, the information given in this section is subject to change; it should be viewed as preliminary.

Introduction

The WFIRST-AFTA will make four major surveys:

- **The High-Latitude Imaging Survey** (HLS Imaging) will enable weak-lensing shape measurements of hundreds of millions of galaxies, which will in turn yield precise measurements of distances and matter clustering through measurements of cosmic shear, galaxy-galaxy lensing, and the abundance and mass profiles of galaxy clusters.
- **The High-Latitude Spectroscopic Survey** (HLS Spectroscopy) will measure redshifts of tens of millions of galaxies via slitless (grism) spectroscopy. Measurements of baryon acoustic oscillations (BAO) in these enormous 3-dimensional maps will pin down the cosmic distance scale and expansion rate at lookback times of 8 – 11 billion years (redshift $z = 1 – 3$), while measurements of redshift space distortions (RSD) will determine the growth rate of matter clustering over the same period.
- **The Supernova Survey** will combine wide-field imaging and Integral Field Unit (IFU) spectroscopy to measure precise distances to thousands of Type Ia supernovae (SNe). IFU observations will also be obtained in parallel with other HLS observations.
- **The Microlensing Survey** of the galactic bulge will search for microlensing events produced by planets in orbit around F, G, K, and M stars and beyond the "snow line", complementing the exoplanets found by Kepler, which orbit closer to the star.

The four surveys will be scheduled according to a draft Design Reference Mission shown in the figure below (source: SDT 2015 Report, p. 139).

Graphical Design Reference Mission

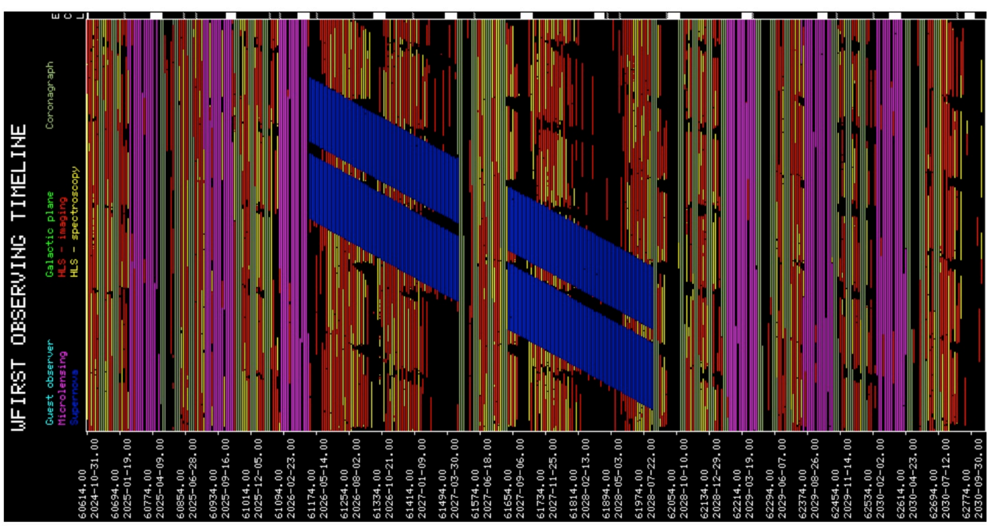


Figure 3-53: The example WFIRST-AFTA observing sequence. Each vertical stripe denotes a 5-day period, and the 6-year mission proceeds from left to right. The microlensing seasons (magenta) and coronagraph blocks (gray) are indicated, as are eclipse seasons (white rectangles at top). The supernova survey (blue) uses portions of the telescope time for 2 years. High-latitude imaging (red) and spectroscopy (yellow) are interspersed.

3.5 Calibrations. The four major surveys described above will not only yield important scientific information directly but will be used to produce essential calibrations. These include:

- Calibration of photometric redshifts with spectroscopic redshifts
- Astrometric fiducials for high-latitude stars
- Characterization of the Wide-Field Imager (WFI) sensor array

Scope of this Report

As described in the 2015 Report, all four surveys will produce important information for cosmic-origins science, *without modification to the observing program*. This section describes the observational data WFIRST will obtain in these surveys and the processing they will undergo *in order to be useful for Cosmic Origins-related studies*. We expect that in most cases the data processing for Cosmic Origins science will coincide with that needed for dark energy and microlensing studies. Nevertheless, this report should not be construed as a prescription for data processing by the survey teams, who know best how to treat the data for their own purposes. This report speaks only on behalf of Cosmic Origins scientists.

Another important caveat is that the WFIRST-AFTA SDT was not charged with developing a plan for data processing, archival and distribution. Nor are the detailed techniques for treating the data a settled matter. For example, it is not clear at this point whether instrumental distortions will be removed before or after galaxy shape measurements proceed.

Data Processing Levels

NASA customarily classifies science data according the level of processing received. The data-processing steps are described in general terms. Here, we describe the data processing steps as applied to WFIRST-AFTA data. Typically (but not always), the WFIRST science data-processing system:

- *Acquires science data* (e,g, images, spectrograms) and engineering data (e.g. temperatures) in the telemetry data and corrects for/removes telemetry artifacts and extraneous data (Level 0 data).

  At this early stage of WFIRST-AFTA, the contents of the science and engineering data down-linked to the ground are uncertain because the orbit of WFIRST-AFTA has not been selected. If the orbit is geo-synchronous then all the data read out from the Wide Field Imager (WFI) every 5 sec or so can be sent to the ground, but if the orbit is about the L2 libration point, only selected data – say, the counts accumulated after 6 non-destructive readouts (30 s) – can be sent to the ground. In the latter case, the data would be processed on-board and corrected for cosmic-ray artifacts according to an algorithm such as Offenberg, Fixsen, & Mather (2004). However, the selected orbit and on-board processing are not settled matters at this time.

- *Gathers individual exposures* into an observation set (multiple observations at a nominal pointing but different dither positions), resamples, and reformats the data into a single image, spectrogram, or datacube in FITS format, and in the process, corrects for cosmic-ray and instrumental artifacts (Level 1 data);

- *Removes the instrument signature* using calibration data and algorithms to yield calibrated data (Level 2 data). Both photometric calibration and geometric calibration (correction for instrumental optical distortions) are carried out. Individual spectra are extracted from the grism spectrogram or IFU datacube, and wavelengths are assigned to each spectral pixel.

- *Operates on and makes measurements on the calibrated data* to produce a Source List with measured properties (Level 3 data). Each observation set produces a Source List.

- *Associates each Source with an Object* in various Object Lists, which contain or have pointers to all measured properties, spectra, postage-stamp images, etc. of each object observed by WFIRST-AFTA (Level 4a data).

- *Combines/resolves differences in the measured properties of an Object* derived from different WFIRST-AFTA observations and possibly other observatories (Level 4b data).

- *Produces "catalogs" of objects* (in reality, on-line relational databases, maps, etc.) of use to the astronomical community (level 4c data). Some examples:

    - Catalog of galaxy spectroscopic redshifts resulting from IFU observations

    - Catalog of galaxy photometric redshifts derived from LSST and WFIRST-AFTA filter fluxes

    - Catalog of proper motions of stars in the galactic bulge derived from numerous astrometric measurements over the course of the observing program

    - Catalog of SNe, with classifications, redshifts, etc. and association with host galaxies

The Object Catalog will be extended as new observations arrive and revised as new calibration algorithms and data become available. Since most of the surveys will take the full primary mission (6 years) to complete, the Object Catalog at any point in time should be considered a work in progress. If WFIRST-AFTA follows the example of SDSS and LSST, the full data set will be reprocessed at least annually. This means full reprocessing of the observed data from Level 1 on.

High-Latitude Imaging Survey (HLS Imaging)

Over its six-year primary mission, the Wide-Field Imager (WFI) will map ~2,200 deg^2 of sky in four broad NIR passbands (Y, J, H, F184) down to a 5-sigma limiting AB magnitude of J=26.9 (point source). The WFI field of view is 0.281 deg^2. The image is recorded by 18 4k x 4k HgCdTe sensor arrays (288 million pixels; ~0.6 Gigabytes). Figure 1-1 (next page) shows its footprint on the sky. Each square pixel subtends 0.11 arcsec. The exposure time for each of the filters is 174 s at each dither position. The HLS imaging survey follows a dithering strategy (Fig 3-54 on next page) so that images in J, H, and F184 are fully sampled for galaxy shape measurements, even when some exposures of a galaxy are lost to chip gaps, cosmic rays, or detector array defects. Thus, a given galaxy is observed 8-10 times in the J filter, 5-8 times in Y, H, and F184. (J-band images get extra dithers for improved PSF sampling.)

The total exposure time for the imaging survey is 2.01 years including overheads. The WFI observations are expected to detect, identify, and measure the shapes of 380 million galaxies. The shape measurements are based on observations in 3 filters, 2 roll orientations per filter, and 3-5 dither positions per orientation. The WFI observations are also expected to lead to the discovery of >10,000 high-redshift (z>8) galaxies.

Processing of Wide-Field Images For Cosmic Origins Science

Data Level	Data Description
Level 0	Single-frame, reformatted, unprocessed science image(s) and ancillary data with telemetry artifacts removed
Level 1	Resampled image(s) derived from level-0 images at other dither positions and corrected for cosmic-ray artifacts
Level 2	Level-1 image with geometric, photometric, and astrometric calibrations applied given the best calibration algorithms & data available at the time
Level 3	Measurements of Level-2 images to produce a Source List with measured properties such as source id, source type, filter ID, AB magnitude, position, size and shape
Level 4	Creation of or addition to the HLS Imaging Object List by association of each Source to an Object

Science Data Products: Indexed Relational Databases & Lists (Examples)

- Stellar proper motions of the Sagittarius stream
- Photometric & astrometric (parallax, proper motion) properties of halo stars
- Catalog of rich galaxy clusters
- Overdense regions in image (star clusters or streams, galaxy clusters)
- Catalog of very high-redshift (z-phot>7) galaxies
- List of astrometric fiducials (barely resolved galaxies) for this pointing

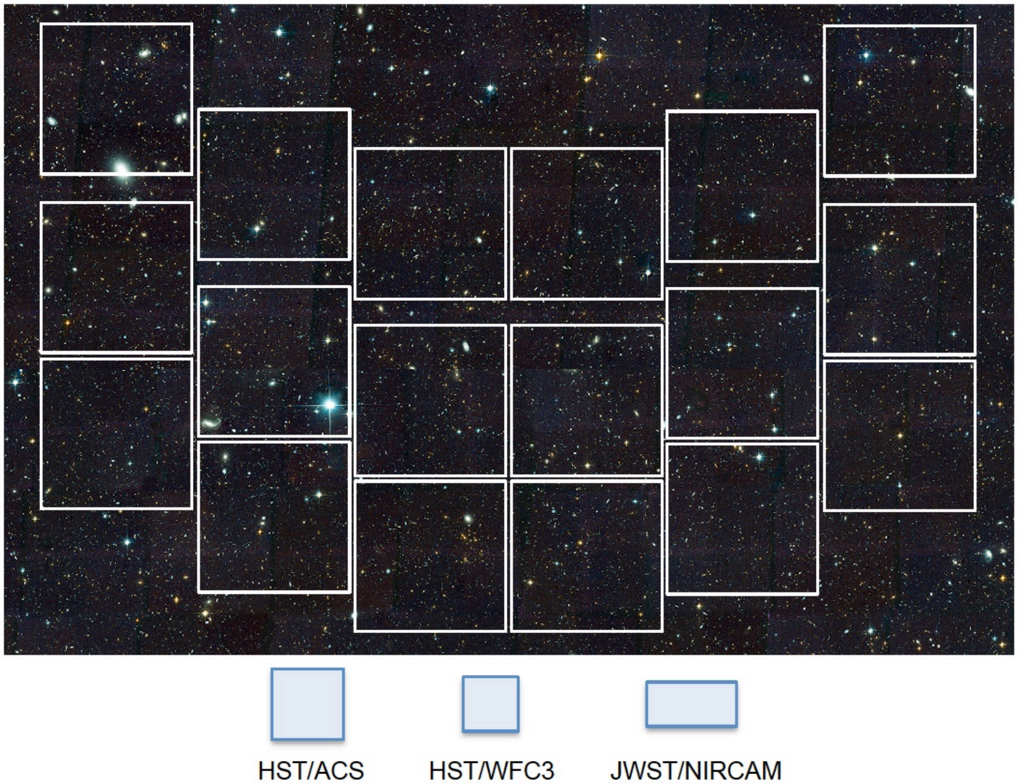


Figure 1-1: Field of view comparison, to scale, of the WFIRST-AFTA wide field instrument with wide field instruments on the Hubble and James Webb Space Telescopes. Each square is a 4k x 4k HgCdTe sensor array. The field of view is 0.28 degrees^2. The pixels are mapped to 0.11 arcseconds on the sky.

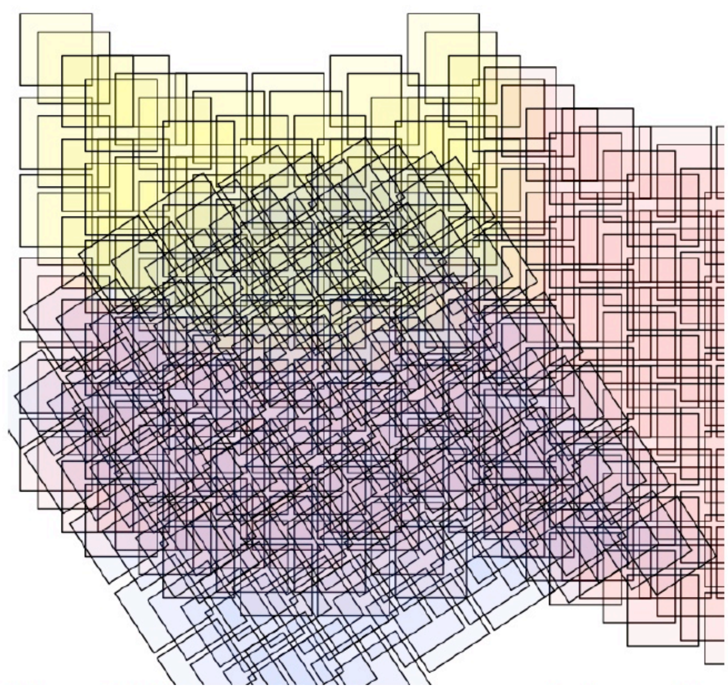


Figure 3-54: Example of the sequence of observations for the high-latitude survey in H-band. The observatory first performs a sequence of 4 small-step dithers to cover the chip gaps (yellow). This pattern is repeated to cover the sky (red). A half-integer number of years later, the observatory returns to perform a second pass over the field (blue) at a general angle. The second pass enables internal relative calibration via repeat observations of stars at all pairs of positions on the focal plane, as well as monitoring of long-term drifts.

High Latitude Spectroscopic Survey

The High Latitude Spectroscopic Survey will make use of a grism to obtain slitless spectrograms over the same region of the sky as the HLS imaging map (Fig. 2-16 below). The exposure time for each observation (dither postion, roll angle) will be 347 s. The observation will be repeated after a small dither, and another pair of pointings will be obtained after a roll of 4-5 degrees. This observation set will be repeated at ≧ 2 different telescope roll angles, two of which are approximately opposed. When co-added together, the R~600 spectra covering 1.35-1.95 micron will have a 7-sigma line-flux sensitivity of 0.5×10^{-16} erg/s/cm^2 (point source). The survey is expected to detect and identify 16 million Hα-emitting galaxies at *z*=1.06-1.88, and 1.4 million [O III]-emitting galaxies at *z*=1.88-2.77. It will also obtain spectra of non-emission-line objects.

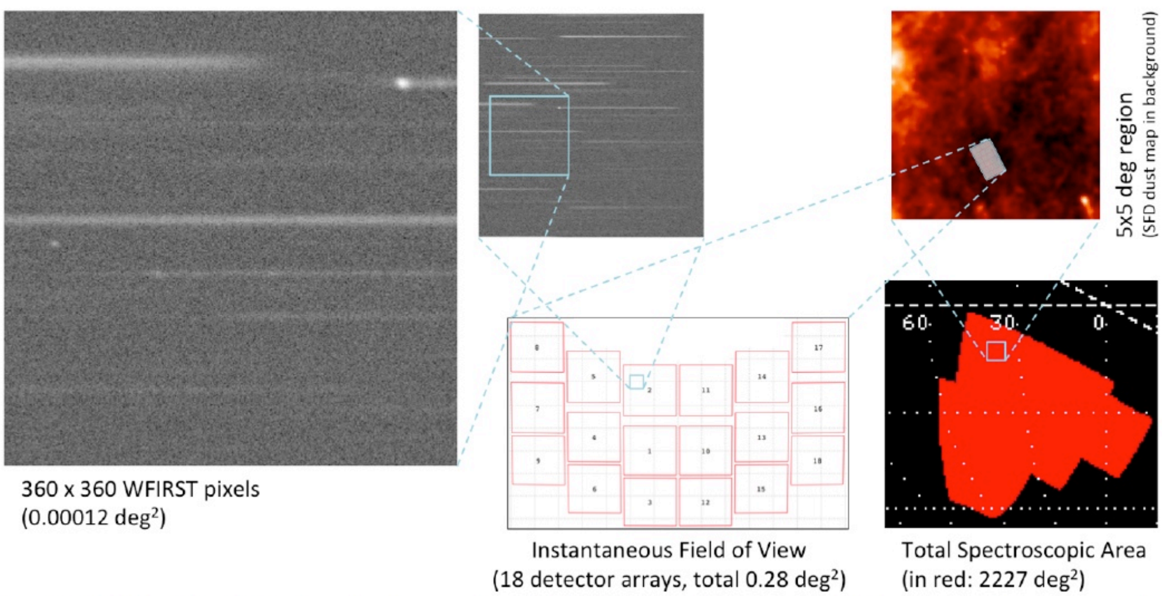


Figure 2-16: A simulated 1st order grism image at WFIRST-AFTA resolution and sensitivity, and at 347 s exposure time as planned for the HLS. Each field in the HLS will be covered with a median of 7 grism exposures at 4 roll angles. This image covers 40x40 arcsec; the instantaneous field of view of WFIRST-AFTA is 2300 times larger.

Processing of Grism Spectrograms for Cosmic Origins Science

The full set of grism data at a particular location on the sky consists of 4 pairs of spectrograms: one pair at the initial roll angle, one pair offset by 4-5 degrees, and two pairs offset from the first by 160^{o} -180^{o}. Thus, the basic unit of grism data is a pair of spectrograms at 2 dither positions but the same telescope roll angle.

Data Level	Data Description
Level 0	Two unprocessed spectrograms and ancillary data with telemetry artifacts removed
Level 1	Spectrogram combined from the two dithers and corrected for cosmic-ray artifacts
Level 2	Extracted spectra of stars and galaxies with wavelength and flux calibrations applied given the best calibration algorithms & data available at the time
Level 3	Creation of or addition to Source List including measured galaxy ID, wavelengths and fluxes (erg/s/cm^2) of identified spectral lines, estimated spectroscopic redshift, and pointers to the calibrated spectra.
Level 4	Co-addition of spectra of each Source at all dithers and telescope roll angles and association of each Source to an Object in the Object List.

Science Data Products: Indexed Relational Databases & Lists (Examples)

- Galaxy Redshift database
- Hα-emitting galaxies - properties
- Luminous red galaxies- properties
- High-redshift galaxies/AGN's - properties

IFU Survey of Supernovae and Galaxies

The Supernova Survey field is a 27.44 deg^2 field (~100 WFI fields) consisting of three nested tiers: a shallow survey for z<0.4 SNe covers the whole field; a medium-deep survey covers 8.96 deg^2 for z<0.8 SNe; and a deep survey covers 5.04 deg^2 for SNe out to z=1.7. This field will be surveyed in just over six months total observing time spread over 2 years. It is expected that 2725 SN Ia's will be discovered.

Before the survey begins in earnest, the field will be observed with multi-filter imaging with the Wide Field Imager (WFI) (and, ideally with spectra from the Wide Field Grism, as well) in order to obtain photometric redshifts (and where possible, spectroscopic redshifts) for most of the galaxies that will later host supernovae.

Promising SN Ia candidates will be observed with the Integral Field Unit (IFU) every 5 days (in the rest frame of the host galaxy). The IFU channel uses an image slicer and spectrograph to provide individual R~100 spectra of each slice covering the 0.6 – 2.0 µm spectral range over two small fields: a 3.00 x 3.15 arcsec field with 0.15" slices, and a 6.00 x 6.30 arcsec field with 0.30" slices. Both fields are imaged onto a single HgCdTe detector. The SNIa candidates will be reobserved to build up light curves at multiple wavelengths.

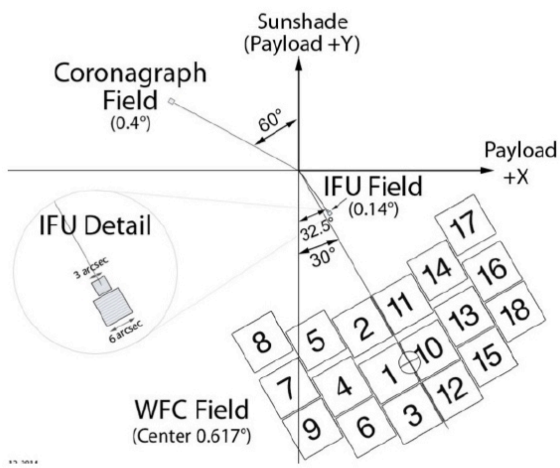


Figure 3-1: The WFIRST-AFTA fields of view layout as projected on the sky showing the wide-field instrument fields (wide-field and IFU channels) and the corona-graph field (shared between the imager and the IFS).

The IFU will also be used to observe in parallel with the Wide-Field Imager (WFI) and Wide-Field Spectrograph (WFS). Keeping step with the WFI or WFS means that the IFU exposure time is also about 173 s and that the IFU will sample multiple small fields for every wide-field observation. As the WFI or WFS returns to its original field (but different orientation), new IFU fields will be observed. In this way, the IFU is expected to obtain spectra covering 0.6-2.0 µm of 20,000 – 30,000 galaxies, thereby providing spectroscopic redshifts for calibrating photometric redshifts.

Processing of IFU Spectra for Cosmic Origins Science

Data Level	Data Description
Level 0	Unprocessed spectrogram with telemetry artifacts removed
Level 1	Level-0 data (spectrogram) corrected for cosmic rays and reformatted into a datacube (x,y,λ), such that a x-y slice through the datacube produces a monochromatic image at a particular wavelength, and a drill through the datacube at a particular x,y yields a R~100 spectrum at that location.
Level 2	Extracted spectra of galaxies in the IFU field of view with wavelength and flux calibrations applied
Level 3	Measurements of the extracted spectra to yield spectroscopic redshifts, emission-line ID's, etc. and creation of or addition to a Source List with pointers to the calibrated spectrum and measured fluxes For SN Ia candidates: Measurement of the SED to yield the flux in selected wavelength bands for that epoch for insertion into a table of light curves for that source.
Level 4	Association of each Source with an Object in an Object list with spectra, spectroscopic redshift, and light-curve fluxes for that epoch (level 4a data). Combination of spectra of each source at different dithers and roll angles (level 4b data) Production of databases containing spectra, spectroscopic redshift, and imaging data. Production of light-curve catalog at all epochs (so far).

Science Data Products: Indexed Relational Databases & Lists (Examples)

IFU spectroscopic redshift catalog
SN Ia Light Curves at various wavelengths
SN Ibc's and other variable sources in the Supernova field
SN Ia Host Galaxies - properties & spectra

Microlensing Survey of the Galactic Bulge

The Wide Field Imager will also be used to make a microlensing survey of the Galactic Bulge. The survey will cover ten contiguous fields between Galactic longitudes ~ -0.4 to 1.8 deg, and Galactic latitudes ~ -1 to 2.2 deg, for a total area of 2.81 deg^2. The ten fields will be imaged in six 72-day campaigns spread over the nominal six-year mission (Fig 3-53 on page 28). In each campaign, each field will be observed every 15 minute in a wide (W149) filter in a 52-sec exposure, and once every 12 hours in a bluer filter (Z087) in a 290-sec exposure. In 85% of the survey footprint, the total number of epochs will be ~33,000 in the W149 filter, and ~700 in the Z087 filter. The strategy on dithers and possible addition of a bluer filter for the Wide Field Imager are under discussion.

Among the 10^8 stars viewed in the microlensing survey, ~2600 bound exoplanets are expected to be detected in the range, 0.1-10,000 Earth masses including ~370 Earth-mass planets, and ~30 free-floating Earth-mass planets if there is one per star in the Galaxy. It will also discover ~20,000 transiting planets with orbits less than a few tenths of an AU.

The microlensing survey is also expected to:

- Measure or constrain the mass function of compact objects over roughly 8 orders of magnitude, from isolated compact objects with masses like that of Mars to roughly 30 times the mass of the Sun;
- Detect and measure the mass of over 100 neutron stars and black holes;
- Yield a very deep color-magnitude diagram of the ~3 deg^2 microlensing field;
- Measure the luminosity function of stars in the bulge down to nearly the bottom of the main sequence, as well as to identify unusual stellar populations;
- Identify many variable sources with amplitudes greater than few mmag, timescales tens of minutes up to the ~6-year duration of the mission;
- Determine the mass and radius of ~1×10^6 giant stars brighter than H_{AB} =14 stars through asteroseismology;
- Obtain photon-noise-limited parallaxes of <10% and proper motion measurements of <0.3% (0.01 mas/yr) for the ~56 million bulge and disk stars with H_{AB}<21.6 in its field of view.

If combined with photometry in a filter bluer than Z087, the microlensing survey will also yield:

- An estimate of the effective temperature, metallicity, age, luminosity, and foreground extinction for all stars;
- A determination of the bulge mass and velocity distribution (including bar structure);
- The stellar density and velocity distribution of the Galactic disk;
- The metallicity and age distribution of the disk and bulge, and the three-dimensional distribution of dust along the line of sight toward the bulge fields.

Processing of WFI Survey of the Galactic Bulge for Cosmic Origins Science

Data Level	Data Description
Level 0	Reformatted, unprocessed wide-field image with telemetry artifacts removed
Level 1	Level-0 image combined with other dither positions and corrected for cosmic-ray artifacts Image difference (image subtracted from level-1 previous image)
Level 2	Level 1 data with geometric, photometric, and astrometric calibrations applied given the best calibration algorithms & data available at the time
Level 3	Creation of or addition to Bulge Source List with measured properties e.g. source id, time, filter ID, AB magnitude, position, etc.
Level 4	Creation of or addition to the Bulge Object List by association of each source to an object observed at a different epoch

Science Data Products: Indexed Relational Databases & Lists (Examples)

- Catalog of micro-lensing events
- Catalog of massive stellar remnants (black holes and neutron stars) in the bulge field
- Catalog of compact objects with masses ranging from Mars-like masses to 30 $M_\odot$
- Monitor of the WFI detector characteristics (see 2015 Report, Section 2.5.7)
- Catalog of stellar populations in the Galactic Bulge including measurements of position, AB magnitudes at W149 and Z087, color (Z087-W149), proper motion, parallax

Special Calibrations (Reprints from the 2015 Report)

Calibration of Photometric Redshifts (2015 Report, §2.2.3.4)

Photometric redshifts for a sample of hundreds of millions of galaxies out to z=3 will be a huge asset to studies of the evolution and clustering of galaxies and AGN… Photometric redshifts (z-phot) of galaxies in the HLS imaging survey will rely on a combination of WFIRST-AFTA data and optical data from the ground, e.g. LSST. The SDT's simulations indicate that the combination of WFIRST+LSST provides excellent performance out to z>3. The estimated rms uncertainty of z-phot is $\sigma_z/(1+z)=0.061$, and estimated fraction of objects with fractional redshift error < 0.04 is 78%.

One strategy for improving the photometric redshifts involves calibration with spectroscopic redshifts obtained with the IFU. The IFU will observe in parallel with the WFI during its weak-lensing survey, sampling numerous fields of galaxies and yielding an estimated ~20,000-30,000 spectroscopic redshifts for these galaxies.

Astrometric Fiducials for Measuring the Proper Motions of High-Latitude Stars (2015 Report §2.3.3)

WFIRST-AFTA will provide access to an enormous number of slightly resolved medium-brightness galaxies, and absolute motions can be measured directly within the field of each detector array. This local approach to measuring absolute motions has recently been used in the optical with HST to measure the absolute motions of the globular clusters, LMC, SMC, dwarf spheroidals and even M31, in addition to individual hyper-velocity and field stars. The strategy involves constructing a template for each galaxy so that a consistent position can be measured for it in exposures taken at different epochs. This template can be convolved with the PSF to account for any variations in focus or changes of the PSF with location on the detector array. The same approach that worked with HST in the optical should work with WFIRST-AFTA in the IR. [Based on prior experience with HST], the SDT expects about 500 reference objects in each WFIRST-AFTA detector array, which can be used to tie down the absolute frame to better than 0.5 mas in each exposure. The HLS 2-year baseline would allow absolute motions to be derived with systematic accuracies of about 125 μas/year. Follow-on GO programs could extend this baseline to 5-years, enabling accuracies of about 50 μas/year. There are about 10 stars in the ~5 square arc-minute FoV of the UDF that can be measured with this accuracy, implying about 18 million halo stars in the high latitude imaging survey.

Wide-Field Imager Detector Characterization (2015 Report §2.5.7)

With its ~33,000 dithered images of ~10^8 point sources with known colors and nearly constant fluxes, the Microlensing Survey of the galactic bulge will likely provide one of the best datasets with which to characterize WFIRST-AFTA's wide-field imager. In particular, we imagine that the bulge survey pipeline will measure the fluxes and positions of all of the stars in the field of view, while simultaneously measuring the response function of each of the WFI pixels, and how this response function varies with time. This time-variable response function will then be used to calibrate and remove detector artifacts such as persistence, inter-pixel capacitance, reciprocity failure, non-linearity, and intrapixel response variations.

SECTION 4. ACCESSING THE WFIRST-AFTA ARCHIVES

This section focuses on the means by which scientists will identify, access, and extract information from the archives. It recommends a query system as the main interface between the user and the archives. To illustrate how this would work, we give examples of queries that return the most relevant information and can be used to discover complex relationships. The queries given here were inspired (not exact quotes) by:

- the 1-page science ideas, printed in Appendix D of the *WFIRST-AFTA SDT 2015 Report, March 10, 2015*

  http://wfirst.gsfc.nasa.gov/science/sdt_public/WFIRST-AFTA_SDT_Report_150310_Final.pdf

- A. Szalay et al.'s 20 Questions for the SDSS Skyserver

  Szalay et al. (2000), *Designing & Mining Multi-TB Ast. Archives,* Proc. ACM SIGMOD 2000, p. 451

- The *WFIRST-AFTA SDT Interim Report, Apr 30, 2014*
  http://wfirst.gsfc.nasa.gov/science/sdt_public/WFIRSTAFTA_SDT_Interim_Report_April_2014.pdf

- J. Kruk (private communication)

The reference for each query is given in shorthand, e.g. (D-12, Tanner) for A. Tanner's 1-page science idea printed in Appendix D, page 12; (Szalay, Q1) for Question #1 of Szalay's 20 Questions for the SDSS Skyserver; (Gould, IR-57) for the *SDT Interim Report*, pg. 57).

Introduction

A single high-latitude imaging observation will record light from nearby objects (e.g. asteroids, brown dwarfs, cool dwarf stars), halo stars of the Milky Way, stellar streams orbiting the Milky Way, nearby galaxies, and distant galaxies. All these objects are a delight to cosmic-origins scientists.

As described in Section 3, data in the WFIRST archive will be stored in one of two formats. Level 0-2 science and calibration images will be stored in "flat" files, presumably in FITS. For example, Wide Field Imager or Spectrograph data will be stored in 18 FITS extensions -- one extension for each of the 18 detectors. Similarly, IFU data cubes may be stored in 20 FITS extensions – one extension for each slice. In the case of spectra extracted from calibrated grism images or IFU data cubes, the data may be stored in FITS structures with each object having 1-D data (wavelength, flux, propagated error, quality flag, etc.) In contrast, most Level 3+ data will be stored in numerous, heavily indexed, relational databases (often called "catalogs" for brevity in this report). These databases contain the results of measurements on the data, source lists, and object lists, etc.

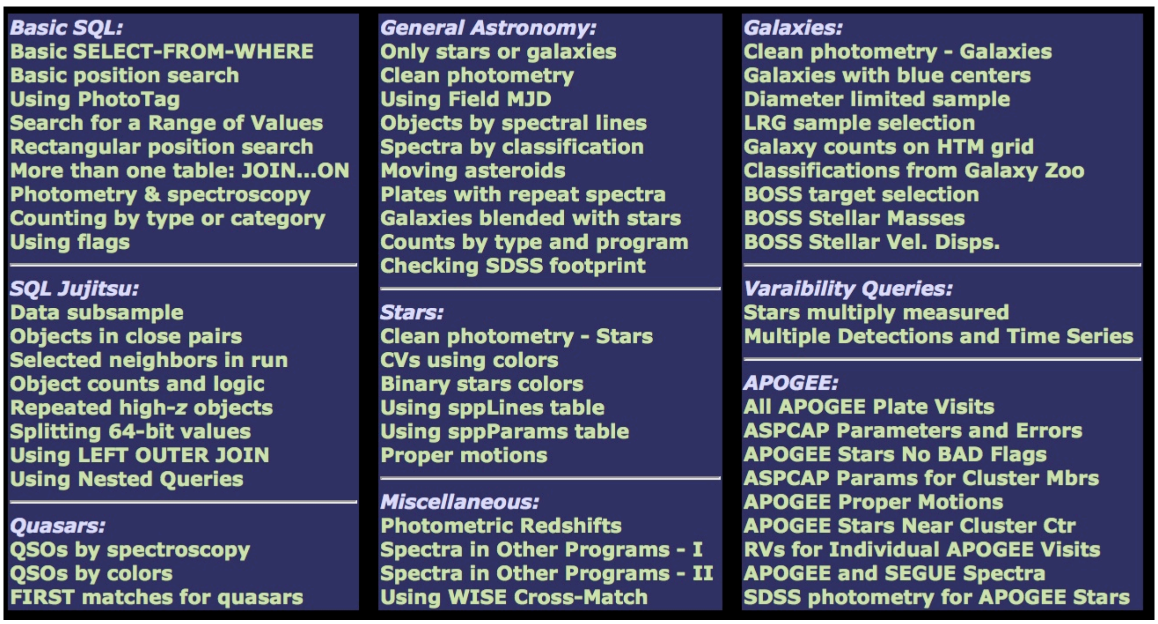
Most data of interest to cosmic-origins scientists will be contained in one or more catalogs rather than the observational-data files from which they were derived. For example, the study of kinematics of stellar streams orbiting the Milky Way (2015 Report §2.3.3 reprinted in Section 2 of this report) will identify stellar streams by their common proper motion as listed in the HLS/WFI Stellar Proper Motion database.

Cosmic-origins scientists will rely on sophisticated Level 3+ data, beyond the most common measurements of position, photometry and line strength. They will expect that the WFIRST archive will contain the output of significant data processing efforts by the science centers and the survey teams. Quantities such as statistical clustering parameters or advanced morphological classifications (galaxy types, lenses, tidal features, etc.) may be required. We expect that advances in computing and software technology, such as machine learning, will enable tremendous advances in what the archive can offer.

Beyond even the advanced data products that the archive can offer, cosmic-origins science will also rely on the capability of the archive to perform new kinds of analysis "near the data". Given the sheer volume of the WFIRST and ancillary data sets, it may be impractical for researchers to do all of the analysis at home. Rather, the archive will likely need to set up the infrastructure to accept (reasonably sized and vetted) software to be run on CPUs with immediate access to the "big data" stored at the archive. This new paradigm will help level the playing field between the few institutions that can afford to build major computing centers and others where good ideas would otherwise be put aside.

Given the complexity of WFIRST (and ancillary) data, providing easy access will be a challenge on its own. Fortunately, Cosmic-Origins scientists don't need to know where the relevant data are stored; they can get the information they need unencumbered by irrelevant data by querying the WFIRST-AFTA archive. The SDSS archives and NASA archives like WISE and 2MASS at IRSA and the Hubble Source Catalog at MAST provide a query system to help a user find the objects of interest. For example, a user might request the database to return the positions of all stars in the Small Magellanic Cloud brighter than some J magnitude and redder than some J-H color.

Most archives use Structured Query Language (SQL) for all queries. These queries can range from the simple to highly complex. The figure below shows examples of queries to the SDSS archives. Users can modify a sample query similar to their own without needing detailed knowledge of SQL. The set of sample queries should be useful for most queries of the WFIRST-AFTA archive.



Source: http://skyserver.sdss.org/dr10/en/help/docs/realquery.aspx

The remainder of this section gives sample questions specific to WFIRST-AFTA observations that can be translated into SQL queries, if the relevant products have been produced by the data processing efforts. These questions can be used to perform stress tests on the WFIRST-AFTA Archive while in development. The questions are inspired (not exact quotes) by the 1-page science ideas (printed in Appendix D of the WFIRST-AFTA 2015 Report), by Szalay et al.'s (2000) *20 Questions for the SDSS Skyserver,* by Gould in an *WFIRST Intermediate Report* (IR), and by J. Kruk.

Sample Queries of the WFIRST Archive

A. MICROLENSING SURVEY (using filters Z087, W149)

ML1: Find all microlensing events of stars in the galactic bulge in which the apparent position of the lens shifted by more than <X > micro-arcsec during the microlensing event. (Sahu, D-18). This is a search for neutron stars and stellar-mass black holes in the Galaxy.

ML2: Provide a list of all bulge stars showing evidence of having a transiting planet(s). (Gould IR-53,57)

ML3: Provide a list of all KBO's detected by WFIRST (Gould 2014, IR-57)

ML4: Find all objects whose absolute magnitudes and colors are consistent with blue stragglers /red giants /young, massive stars/ white dwarfs / <keyword>. (31Oct14 WFIRST SDT telecon)

B. HIGH-LATITUDE IMAGING SURVEY (using filters Z087, Y, J, H, F184)

Milky Way and Local Group Science

HLIL-1: Find all stars brighter than J~25 whose LSST + WFIRST + WISE colors are consistent with an L or T brown dwarf. (D-12, Tanner)

HLIL-2: Find statiscally overdense regions of resolved stars consistent with dwarf-galaxy satellites of the Milky Way. (D-22, Geha)

HLIL-3: Find all stars in a selected nearby galaxy (d<5 Mpc) with absolute magnitudes and colors consistent with red-giant stars. (D-27, van der Marel)

HLIL-4: Find all galaxies within 50 Mpc showing evidence of tidal stellar streams or intracluster light. (D-29, Laine)

HLIL-5: Provide a list of stars that are 1% rare for LSST+WFIRST-color attributes. (Szalay, Q7)

HLIL-6: Provide registered 4-filter images of a selected star-forming region that can be used to create a dust map of the region (IR-71)

HLIL-7: Provide a map of absolute proper motions of stars in the Small Magellanic Cloud (2015 Report §2.3.3).

Science Beyond the Local Group

HLIG-1: Find all galaxies within <x> arcsec of a given point in the sky (Szalay, Q1). This is a classic spatial lookup.

HLIG-2: Find all objects within 1' of one another that have similar colors. This is a gravitational lens query. (Szalay, Q18)

HLIG-3: Find all objects with rings or arcs. This is a search for strongly lensed galaxies (Stern, D-33; Fan, D-43; IR-48)

HLIG-4: Find all lensed galaxies images brighter than J=<X> (i.e. bright enough for follow-up spectroscopy by large ground-based telescopes).

HLIG-5: Find all galaxies with central surface brightness fainter than J = 23 mag per square arcsecond. (Szalay, Q2)

HLIG-6: Find all galaxies showing a tidal feature with no obvious companion in the H band (Conselice, D-32)

HLIG-:7 Find galaxies that are blended with a star; output the deblended magnitudes. (Szalay, Q6)

HLIG-8: For a selected massive galaxy cluster, find all galaxies within 30" of the brightest cluster galaxy (BCG) with a photo-z within 0.05 of the BCG. (Szalay, Q20)

HLIG-9: Find all galaxies satisfying user-specified color cuts and magnitude & redshift ranges (Szalay, Q13).

HLIG-10: *Find all z>8 galaxies redder than (H-F184)=<X>*

HLIG-11: Find all regions of the sky where the surface density of galaxies observed in the H-band is greater than 50 galaxies per square arc minute (search for galaxy clusters, 2015 Report §2.3.2)

HLIG-12: Find all galaxies brighter than J=26 with shape measurements and 4-filter z-phot>7.5 (search for lensed high-z galaxies, 2015 Report §2.3.1)

C. HIGH LATITUDE SPECTROSCOPIC SURVEY (λ range=1.35-1.95 μm)

NOTE: Some of these queries require data from not only WFIRST (YJHF184) imaging and spectroscopic catalogs (λλ 1.35-1.95 mm) but also the LSST (*ugrizy*) catalogs and the Euclid spectroscopic catalog (λλ 1.25-1.85 mm).

HLS-1: Provide a histogram of H-alpha fluxes and equivalent widths as a function of spectroscopic redshift in the interval, z~1-2. Do the same for [O II] 3727 in the redshift interval, 2.6-4.3. (2015 Report §2.3.4)

HLS-2: Provide a plot showing IFU z-spec's vs. LSST+WFIRST z-phot's (Report 2.2.3.4)

HLS-3: Provide a color-luminosity plot of weak-lensing galaxies in a selected narrow redshift bin with symbols/contours indicating the size of (z-phot-z-spec) error (Report 2.2.3.4)

HLS-4: Find spectra of all Lyman alpha emitters (LAE's) at z>7.

HLS-5: Find all quasars at z=8-10 (D-44, Fan)

HLS-6: Provide spectroscopic redshifts for LSST objects of interest identified by their RA, Dec, and object type. (D-48, Strauss)

HLS-7: Provide IR SED's for LSST stars of interest (D-48, Strauss)

HLS-8: Find all elliptical galaxies at z>1 whose spectra show an anomalous emission line (Szalay, Q11)

HLS-9: Find all objects with spectra unclassified. (Szalay, Q8)

HLS-10: Find quasars with a broad absorption line in their spectra and at least one galaxy within 10". (Szalay, Q19)

*HLS-11: Find quasars with a line width >2000 km/s and 2.5<redshift<2.7***.** *(Szalay, Q9)*

HLS-12: Find galaxies with spectra that have an equivalent width in Hα >40Å (Szalay, Q10)

E. TIME-DOMAIN INFORMATION

TD-1: Find all the objects in the microlensing survey exhibiting peaks in their lightcurve with significance greater than 5-sigma and with durations shorter than 10 days. (Kruk)

TD-2: Download all observations encompassing a given spot on the sky within ±<X> days of a given date. (Kruk)

TD-3: Find all objects whose J-band fluxes have changed by more than <X%> since measured by HST. (Kruk)

TD-4: Find all galaxies in clusters at z~0.5-1.5 in which >0.5-mag variability was detected at F184W (Kruk)

F. IFU SURVEY

IFU-1: Provide a list of all red galaxies showing Hα in emission.

IFU-2: Download the light curves, redshifts, and imaging data of all SN Ia's

IFU-3: Provide the AB magnitudes, sizes and shapes of all host galaxies of SN Ia's

IFU-4: Provide a list of galaxies showing variability not associated with an SN Ia

IFU-5: Provide postage-stamp images of all galaxies hosting a SN Ia, both with the bright SN and after the SN has faded

G. GENERAL OBSERVER OBSERVATIONS

GO-1: Provide a list of globular clusters / nearby galaxies / galaxy clusters / <keyword> observed by WFIRST

GO-2: Download the complete Object List containing all associated measurements for the WFIRST Deep Field